XXVII Fonetiikan päivät

Phonetics Symposium 2012

Tallinn, Estonia

February 17-18, 2012

XXVII Fonetiikan päivät – Phonetics Symposium 2012

Tallinn, Estonia, February 17-18, 2012

Phonetics Symposium 2012 (XXVII Fonetiikan päivät) continues the tradition of meetings of Finnish phoneticians started in 1971 in Turku. These meetings, held in turn at different universities in Finland, have been frequently attended by Estonian phoneticians as well. In 1998 the meeting was held in Pärnu, Estonia, and in 2012 it will take place in Estonia for the second time.

Phonetics Symposium 2012 (XXVII Fonetiikan päivät) will provide a forum for scientists and students in phonetics and speech technology to present and discuss recent research and development in spoken language communication.

In the recent years we have lost three world-renowned scholars in the phonetic sciences – **Ilse Lehiste**, **Matti Karjalainen** and **Arvo Eek**. The symposium shall commemorate and honor their scientific contributions to Estonian and Finnish phonetics and speech technology.

The symposium is hosted by the Institute of Cybernetics at Tallinn University of Technology (IoC) and organized in co-operation with the Estonian Centre of Excellence in Computer Science, EXCS (funded mainly by the European Regional Development Fund).

XXVII Fonetiikan päivät – Phonetics Symposium 2012

Tallinn, Estonia, February 17-18, 2012

SCIENTIFIC PROGRAM: DAY 1

9:45	**Opening**	
SESSION 1: Chair Karl Pajusalu		
10:00	Unto K. Laine	Professori Matti Karjalainen, suomalaisen puheteknologian edelläkävijä
10:30	Diana Krull	Remembering Arvo Eek and Ilse Lehiste
11:00	Pire Teras	Arvo Eek and Estonian Phonetics
11:30	**Coffee**	
SESSION 2: Chair Kari Suomi		
12:00	Martti Vainio, Antti Suni, Anja Arnhold, Tuomo Raitio, Henri Seijo, Juhani Järvikivi, Paavo Alku	Effect of level and type of noise on focus related prosody
12:30	Osmo Eerola, Janne Savela	Production of short and long Finnish vowels with and without noise masking
13:00	Michael O'Dell, Tommi Nieminen, Mietta Lennes	Modeling turn-taking rhythms with coupled oscillators
13:30-14:30	**Lunch**	
SESSION 3: Chair Paavo Alku		
14:30	Tanel Alumäe	Recent developments of Estonian speech recognition applications
15:00	Tuomo Raitio, Antti Suni, Martti Vainio, Paavo Alku	Recent developments of statistical parametric speech synthesis system GlottHMM
15:30	Tomi Kinnunen, Zhi-Zheng Wu, Kong Aik Lee, Eng Siong Chng, Haizhou Li	Vulnerability of speaker verification systems against voice conversion spoofing attacks
16:00	Stina Ojala, Pertti Palo	A sign annotation tool for phoneticians
16:30	**Coffee**	

SESSION 4: POSTERS		
17:00-18:00	Terhi Peltola	Factors affecting the categorizing and rating the goodness of L2 vowels
	Henna Tamminen, Maija S. Peltola, Katri Jähi, Teija Kujala, Heikki Hämäläinen, Risto Näätänen	Ikääntyminen ja vieraan kielen havaitseminen
	Kimmo Peltola, Antti Saloranta, Henna Tamminen, ja Maija S. Peltola	Ääntämisharjoitusten vaikutus vieraan kielen oppimiseen
	Teija Waaramaa, Mona Lehtinen	Emotions in cents: A preliminary study of the musical intervals in emotional speech samples
	Kairi Tamuri, Meelis Mihkla	Emotsioonid ja kõne ajaline struktuur
	Sulev Iva	Pilguheit võru häälikusüsteemi
	Eva Liina Asu and Nele Salveste	The phonetic and phonological analysis of the fall-rise intonation pattern in the Kihnu variety of Estonian
	Dennis Estill	Revisiting the Meadow Mari vocalic system
	Cemal Hanilci, Tomi Kinnunen, Rahim Saeidi, Jouni Pohjalainen, Paavo Alku, Figen Ertas	Regularized all-pole models: experiments with text-independent speaker verification in noisy environments
	Ville Hautamäki, Kong Aik Lee, Anthony Larcher, Tomi Kinnunen, Bin Ma and Haizhou Li	Sparse logistic regression fusion for speaker verification
	Stina Ojala, Einar Meister	Comparison of two data collecting systems for sign language articulation studies
	Einar Meister, Lya Meister, Rainer Metsvahi	New speech corpora at IoC
19:00-20:30	**Conference dinner**	

SCIENTIFIC PROGRAM: DAY 2

SESSION 5: Chair Einar Meister		
9:30	Karl Pajusalu	Finno-Ugric prosody project
10:00	Niina Aasmäe, Karl Pajusalu, Nele Salveste, Tatiana Zirnask	Quantity in Moksha-Mordvin
10:30	Helen Türk, Eva Liina Asu, Pärtel Lippus, Ellen Niit	The acoustic characteristics of monophthongs and diphthongs in the Kihnu variety of Estonian
11:00	**Coffee**	
SESSION 6: Chair Reijo Aulanko		
11:30	Kari Suomi, Einar Meister	A preliminary comparison of Estonian and Finnish stops
12:00	Mikael Hyvärinen, Pauli Härkönen, Michael O'Dell	Tavunalkuisten konsonanttiyhtymien ajoitus suomessa
12:30	Pertti Palo, Daniel Aalto, Olli Aaltonen, Risto-Pekka Happonen, Jarmo Malinen, Jani Saunavaara, and Martti Vainio	Articulating Finnish vowels: Results from MRI and sound data
13:00-14:00	**Lunch**	
SESSION 7: Chair Maija Peltola		
14:00	Stefan Werner, Einar Meister	Microduration in Finnish and Estonian vowels revisited: methodological musings
14:30	Heikki Rasilo, Unto K. Laine, Okko Räsänen	Virtuaalilapsi oppii puhumaan vanhemman antaman palautteen ja suorittaman matkimisen avulla
15:00	Henna Tamminen, Maija S. Peltola, Teija Kujala, Risto Näätänen	Uudet muistijäljet syntyvät jo kolmessa päivässä
15:30	**Coffee**	

SESSION 8: Chair Stefan Werner		
16:00	Maija S. Peltola, Henna Tamminen, Teija Kujala, Risto Näätänen	Harjoittelu johtaa pysyviin neuraalisiin muutoksiin – uusien muistijälkien juurtuminen vie aikaa
16:30	Riikka Ullakonoja	Suomen vokaalin kesto venäjää äidinkielenään puhuvilla alkeisoppijoilla
17:00	Kati Järvinen, Anne-Maria Laukkanen	Native Finnish and English speakers' fundamental frequency, sound pressure level and long-time average spectrum characteristics in text-reading in Finnish and English
17:30	**Closing**	

Professori Matti Karjalainen, suomalaisen puheteknologian edelläkävijä

Unto K.Laine
Aalto-yliopisto, Espoo

Matti Karjalainen syntyi 2. huhtikuussa 1946 pienviljelijäperheeseen Hankasalmella. Matti kävi lukion Lievestuoreella ja asui viikot mm. perheessä, missä talon isäntä oli radioamatööri. Lukiolaisten juhlissa Matti huolehti usein äänentoistosta. Myös kitaransoitto alkoi kiinnostaa. 50-luvun lopulla alkoi Suomessa saada harrastelukäyttöön transistoreita ja kirjoja, joista löytyi elektroniikan perusteita ja rakennusohjeita. Niinpä jo lukiolaisena Matti alkoi rakennella radiolaitteita. Ylioppilaskirjoitusten jälkeen hän pyrki Teknilliseen korkeakouluun. Samana syksynä perustettiin Tampereelle sivukorkeakoulu ja koska se oli lähempänä Hankasalmea, Matti päättikin aloittaa opintonsa Tampereella.

Samat harrastukset jatkuivat opintojen ohella Tampereen teknillisessä korkeakoulussa, joka sijaitsi tuolloin Pyynikillä. Matti mm. suunnitteli ja toteutti teekkareiden tempauksia varten kuljetettavat, tehokkaat äänentoistolaitteet. Huippuluokan äänentoistolaitteet syntyivät myös itselle ja ystäville vuosien 1970-72 aikana. Diplomityö valmistui 1970 ja lisensiaatin työ, joka käsitteli tehoelektroniikkaa, neljä vuotta myöhemmin. Myös musiikin kuuntelu kiinnosti ja kesästä 1971 lähtien kohteina olivat usein Pori Jazz ja Ruisrock.

Ahkeralle ja lahjakkaalle opiskelijalle avautui korkeakoululla pian assistentin tehtäviä ja yliassistentin toimi. Elokuussa 1973 eräs sattuma muutti Matin työuran suuntaa pysyvästi. Kyseessä oli yhden viikonlopun aikana tehty elektronisen musiikin kokeilu laboratorion tiloissa. Tämän yhteydessä syntyi säätyvä resonaattorikytkentä, joka alkoi yllättäen tuottaa puhetta muistuttavia efektejä. Katalysaattorina ja innoittajana näihin kokeiluihin toimi allekirjoittanut. Heräsi kysymys, voisiko tältä pohjalta rakentaa puhuvan koneen. Sairaalaelektroniikkaan erikoistunutta tutkimusyksikköämme johti tuolloin Professori Boris Segerståhl, joka piti vammaisten apuvälineiden kehitystyötä tärkeänä. Yhteisten pohdintojen jälkeen puhesynteesiä päätettiin ryhtyä tutkimaan ja kehittämään mm. vammaisten kommunikaatiotarpeisiin. Syksyn -73 aikana laadittiin hakemus Suomen Akatemiaan ja niin keväällä 1974 alkoi Tampereella suomenkielisen puhesyntetisaattorin kehitystyö, minkä ensimmäisiä tuloksia esiteltiin jo IV Fonetiikan päivillä Tampereella huhtikuussa 1974.

Puhesynteesin tutkimuslinjan avaaminen oli merkittävä päätös ajatellen koko puheteknologian tutkimuksen ja myöhemmin opetuksenkin kehittymistä Suomessa. Matti Karjalainen väitteli aihepiiristä vuonna 1978 ja hänestä tuli akustiikan apulaisprofessori TKK:lle 1980. TKK:lla Matin perustamien puheteknologiaan liittyvien kurssien kautta on kulkenut satoja opiskelijoita.

Seuraavien vuosikymmenten aikana Matti laajensi tutkimustyötään uusille sektoreille, kuten psykoakustinen mallinnus, symbolinen tiedon ja signaalien käsittely, puheentunnistus, keinotekoiset hermoverkot, puhetietokannat, 3D-audio (tilakuuleminen) ja akustisten soittimien mallinnus. Hänen tutkimussarkansa monipuolisuus ja tieteellinen tuottavuus hakee vertaansa niin kansallisesti kuin kansainvälisestikin. Tästä kertovat myös monet kansainvälisten tiede- ja tutkimusorganisaatioiden taholta tulleet huomionosoitukset. Viimeiset vuodet Matti hoiti professuuriaan sitkeästi vaikeasta sairaudestaan huolimatta. Matti Karjalainen kuoli kotonaan Espoossa 30.5.2010.

Remembering Arvo Eek and Ilse Lehiste

Diana Krull

Stockholm University, Deptartment of Linguistics

Arvo Eek

After Estonia became independent again in 1991, new possibilities opened up for Estonian scholars to work abroad. In Sweden, funds became available for inviting scholars to work as visiting researchers and create contacts with Swedish colleagues. I myself made use of this possibility. Similarly, many other Estonians were able to come to Sweden and work here for short periods. One of them was Arvo Eek. He was already known among many Swedish phoneticians through the ICPhS XI held in Tallinn 1987 of which he helped organize. Many of us had also read his articles in Estonian Papers of Phonetics and appreciated his work.

Arvo Eek was in Stockholm several times. He appears on the list of visitors to the phonetics laboratory at Stockholm University during 1992, 1993 and 1994. Here he could use modern technology not yet available in Estonia, and this was of course of great use to him in his studies of the acoustics of Estonian vowels. However, I myself probably benefitted even more from his stay here. I started my studies in phonetic late in life and there was much in Estonian phonetics that was new to me. We had long talks about Estonian quantities, a subject on which he was very knowledgeable. I also used him as a speaker for my studies of Estonian spontaneous speech. He could easily keep up a lively monologue for almost an hour, telling stories from his childhood and from his travels in the Soviet Union from Yakutsk to Tajikistan in connection with his work with the maintenance of native languages within the Union.

In the following years we continued writing to each other. The last letter came in the beginning of December 2008 where he wrote that wanted to wish me a happy Christmas so early because he had to undergo a difficult operation, and was not sure if it would be possible for him to do it later.

Ilse Lehiste

I first met Ilse Lehiste in the beginning of the 90-ies. In the following years we met mostly at phonetics conferences in USA and Europe, twice in Estonia, and I also visited her at her home in Columbus, Ohio. However, most of our contacts were through the e-mail. I had read many of her papers and liked them very much. She also read my papers and commented on them. Our exchanges concerned Estonian phonetics in particular. We had had a common interest – the acoustics and perception of the three quantity degrees that are special to Estonian. She had worked in this field much longer than I and could give me useful advice, directing my attention to details I had not noticed. She also gave me much encouragement in my work. For my part, I was glad to be able to help her with a perception test involving Swedish listeners and with arranging a part of the financial support for a phonetics conference in Tallinn of which she was one of the organizers.

I received advice from her also in other matters, e.g. concerning helping young phoneticians in Estonia. We also discussed the many changes in Estonian research institutions that were taking place after the collapse of the Soviet regime.

We were not always of the same opinion about everything. In particular, our views on the perception of quantity differed significantly. Together with Hartmut Traunmüller – who is a specialist on speech perception – I wrote a paper on the perception Estonian quantity where we proceeded from the assumption that the perception of quantity is segmental. Ilse Lehiste saw it as hierarchical, i.e. not only between segments, but also between higher units. When she was asked by Phonetica to review our paper, she refused, saying that she did not feel impartial concerning the views expressed in it.

Ilse Lehiste and I discussed not only work problems, we also wrote about private matters and sent birthday and Christmas greetings. Among the first Christmas greetings every year was card with a picture of polar bears with good wishes from Ilse Lehiste. In December 2009 there was no such card. I missed it and I still miss Ilse Lehiste.

Arvo Eek and Estonian Phonetics

Pire Teras

University of Helsinki, University of Tartu

One of the greatest Estonian phoneticians Arvo Eek was born in 1937 and died in 2009. His interest in phonetics developed probably during the university time. As an undergraduate student he wrote his first pieces of research on phonetics. In his seminar paper written in 1961 he studied the vowels of Varbla sub-dialect of Western Estonian dialect. His diploma thesis written in 1963 dealt with the problems related to the pronunciation and perception of Estonian quantity degrees. This work combined the two sides of Arvo: pedagogue and researcher. In the same year Arvo became a graduate student at the Institute of Language and Literature of the Estonian Academy of Sciences in Tallinn. In 1971 he defended his Candidate of Philology thesis "Articulation of the Estonian Sonorant Consonants". He became a head of the phonetics laboratory at the Institute of Language and Literature and later he worked there as a senior researcher. Arvo was an editor-in-chief of the Estonian Papers in Phonetics during the whole time it was published (1972–1986). During this time he was interested both in segmental phonetics (acoustic and articulation of Estonian sonorant consonants (including palatalised consonants), nasalisation), and suprasegmental phonetics (acoustics and perception of Estonian quantity degrees). In his works he used methods of articulatory, acoustic and perceptual phonetics. During the times when Estonia regained its independence Arvo Eek engaged actively in politics but in the 1990s he returned to phonetics. From 1993 the laboratory of phonetics and speech technology of the Institute of Cybernetics at the University of Technology became his workplace for many years. In 1994 Arvo Eek defended his PhD thesis "Studies on Quantity and Stress in Estonian".

I don't know exactly when Arvo started to plan writing a book on Estonian phonetics but I remember that I heard about the book when I met him for the first time. It was during the Phonetics Symposium – Fonetiikan Päivät held in Pärnu in 1998. I was then a graduate student and I did not imagine that one day I will be the editor of this book. In my paper I would like to speak about the first volume of the book "Estonian Phonetics" ("Eesti keele foneetika I", 2008) and my experience of collaborating with Arvo. I will also speak about the manuscript of the second volume what Arvo worked on intensively before his death but which unfortunately he could not complete.

Effect of level and type of noise on focus related prosody

Martti Vainio1, Antti Suni1, Anja Arnhold1, Tuomo Raitio2, Henri Seijo2, Juhani Järvikivi3, and Paavo Alku2

1 Institute of Behavioural Sciences, (Phonetics and Speech Synthesis Research Group), University of Helsinki, Finland

2 Department of Signal Processing and Acoustics, Aalto University, Finland

3 Language Acquisition and Language Processing Lab, Norwegian University of Science and Technology, Norway

Speakers automatically raise their voice when forced to speak in environmental noise or when the normal feedback mechanism is disturbed. Raising one's voice consists of various physiological means that have different consequences to the phonetic realisation speech. Typically the speakers' f0 is higher – probably as a consequence of added sub-glottal pressure – and the mode of vocal fold vibration is more pressed in order to decrease the slope of the glottal voicesource spectrum. The adaptation of speech to noise in order to increase the signal-to-noise ratio is called the Lombard effect or Lombard reflex to illustrate its involuntary nature. However, the knowledge regarding Lombard speech is fairly general in nature and not very much is known how the reflex influences prosodic changes that are due to specific communicative needs such as, for instance, signaling linguistic focus.

We recorded 25 speakers producing utterances with different information focus conditions in three types of noise with four noise levels. The purpose of the study was to see whether speakers vary their means of producing prosodic focus as a function of both noise level and type. The analysed utterances were replies to three types of questions designed to elicit either a broad focus or two types of narrow focus in simple three word utterances. The typical prosodic patterns for the three focus conditions are well-known for Finnish, which allows us to compare Lombard speech to normal speech in a controlled manner.

Preliminary analysis of the data shows that, regardless of the increase in f0, the typical intonation contours are still produced. Also, and as can be expected, the f0 is raised as a function of noise level. However, the intonation contour is not compressed as the f0 gets higher. The typical utterance final creaky voice is also not as prevalent in Lombard speech as it is normally and disappears altogether in severely noisy conditions.

In our talk we will present a more thorough analysis of our material with regard to both traditional prosodic parameters, as well as the parameters analysed from the glottal voice source.

Production of short and long Finnish vowels with and without noise masking

Osmo Eerola
Faculty of Telecommunication and e-Business, Turku University of Applied Sciences; Centre for Cognitive Neuroscience, University of Turku, Turku, Finland

Janne Savela
Department of Information Technology, University of Turku, Turku, Finland

In order to further examine the possible quality differences between produced short and long Finnish vowels, we studied the f0, formant frequencies F1-F4, and duration of the eight Finnish vowels /a/, /e/, /i/, /o/, /u/, /y/, /æ/, and /ø/ when uttered in carrier words (e.g. /tili/ - /tiili/) in two different masking conditions and without a noise mask. Babble noise at 92 dB SPL was used to simulate a loud, crowded cocktail party, and pink noise at 83 dB SPL an environment with maximum allowed noise level for continuous working.

It was assumed that noisy surroundings may cause hyperarticulation, and possibly accentuate the reported minor quality differences between short and long vowels (Wiik, 1965; Kukkonen, 1990; Iivonen & Laukkanen, 1993; Iivonen & Harnud, 2005; Eerola & Savela, 2011).

Subjects were asked to utter each word five times successively using their normal speech style, first without the noise mask, and then in the masking conditions. The recordings were carried out in an acoustically dampened room, and saved as sound files for later analysis. Praat SW was used both for the recordings and analysis. The noise masks were played through headphones. The sound samples were automatically analyzed using a text grid, in which the steady-state part of each target vowel was windowed. The f0, formants F1-F4, and durations were analyzed by using the Burg method and averaged for individual results.

The detailed differences between short and long vowels, and between masking and non-masking conditions, will be reported and compared to earlier published research. Preliminary results confirm the earlier findings that there are minor differences between short and long vowels. The mean of individual distances in the F1-F2 plane between long and short vowels without noise mask was 52 mel over all vowel categories. Variation was found between vowels: the /e/, /y/ and /ø/ categories had distances of the order of 30 mel, whereas /o/ and /u/ showed clearly larger distances, up to 80 mel.

The effect of the noise was similar in both masking conditions, and no differences between babble and pink noise was found. Both noise types seem to cause higher F1 and F2 frequencies in the production of the mid-high vowels: On the average, the formant values of the short and long vowels produced in the masking conditions are about 50 mel (in Euclidean distances) higher than without masks. No similar effect was found on the low vowels /a/ and /æ/. The largest individual F2 difference between masking and non-masked condition was 70 mel for /u/. The preliminary results indicate that noise masks cause a systematic shift of F1-F2 values in the production of mid-high Finnish vowels.

References

Eerola, O., Savela, J. (2011). Differences in Finnish front vowel production and weighted perceptual prototypes in the F1-F2 space. *Proceedings of the 17th International Congress of Phonetics Sciences*, University of Hong Kong, Hong Kong, China.

Iivonen, A., & Harnud, H. (2005). Acoustical comparison of the monophtong systems in Finnish, Mongolian, and Udmurt. *Journal of the International Phonetic Association*, *35*(1), 59-71.

Iivonen, A., & Laukkanen, A. (1993). Explanations for the qualitative variation of Finnish vowels. *Studies in Logopedics and Phonetics, 4*, 29-55.

Kukkonen, P. (1990). Patterns of phonological disturbances in adult aphasia. Faculty of Arts, University of Helsinki. *Suomalaisen Kirjallisuuden Seuran Toimituksia*, (529), 1-231.

Wiik, K. (1965). Finnish and English vowels. (Doctoral Thesis, University of Turku). *Annales Universitatis Turkuensis, Series B* (94)

Modeling turn-taking rhythms with coupled oscillators

Michael O'Dell, Tommi Nieminen

University of Tampere, University of Eastern Finland

Mietta Lennes

University of Helsinki

Wilson & Wilson (2006) have suggested a model of conversational turn-taking that is loosely based on coupled oscillators. More recently Benuš (2009) tested several predictions of this model against a database of conversational American English. His results provided some support for the model, but the support was weak due to small (although significant) correlations, and a lack of predicted phase patterns.

As Wilson & Wilson pointed out, it is important to gather data on a variety of languages in addition to English. We examine pausing and turn-taking behavior in a Finnish conversational database using an analysis similar to that of Benuš. In addition, we integrate the Wilson & Wilson model into our own coupled oscillator model of speech timing (O'Dell et al. 1998–2010), which has hitherto lacked an explicit mechanism for dealing with pausing behavior. Unlike the Wilson & Wilson model, our model incorporates simultaneous speech rhythms (oscillators) at several hierarchical levels. This enables a finer-grained approach using Bayesian estimation of the model's dynamic parameters.

We expect that the increased flexibility of the model will yield a more detailed picture of synchronization between speakers and a more powerful test of oscillatory behavior than the coarser-grained approach of Wilson & Wilson and Benuš.

Recent developments of Estonian speech recognition applications

Tanel Alumäe

Institute of Cybernetics
Tallinn University of Technology

In the recent year, several Estonian speech recognition applications developed by the Laboratory of Phonetics and Speech Technology at the Institute of Cybernetics at Tallinn University of Technology have been made available to the public. This paper discusses the motivation behind these developments, covers the underlying technology, and introduces the applications.

Language and speech technology research in Estonia is partly funded by the national program *Estonian Language Technology (2011-2017)* which supports projects in many areas and especially encourages creation of research and end-user software prototypes. As a result, we have made efforts in the recent years to make Estonian speech recognition technology developed in our lab accessible to the public via several free applications.

Our core large vocabulary speech recognition technology can be divided into two distinct branches: offline transcription and online recognition. The offline branch aims to achieve high recognition accuracy and does not impose any rigid processing time constraints on itself. It is intended for fully automatic transcription of arbitrarily long speech recordings and incorporates speech/non-speech detection, speaker diarization and uses a multi-pass recognition strategy. In the first pass, speaker-independent acoustic models are applied; the first-pass recognition hypotheses are used for creating a CMLLR feature transform for each speaker, which are used in the second pass of recognition. The second intermediate hypotheses are used for adapting the acoustic models for CMLLR-transformed speech using MLLR. Both transforms are then used to produce the final hypotheses. Such unsupervised adaptation helps to reduce recognition errors relatively by about 15%. The transcription pipeline runs in around six times slower than real-time, relative to the duration of the speech recording. The recognition accuracy has been measured on three types of data: on broadcast news the average word error rate is about 14%, on broadcast conversations (talk shows from radio) about 27% and on a set conference presentation recordings about 34%.

Our online recognition system is intended for recognition of short speech utterances (up to 20 seconds) and aims to deliver recognition result almost in real-time (1.0-1.5 times slower). The system is optimized for real-time processing: decoding is started as soon as the first frames of input speech are received. The word error rate on a set of utterances randomly sampled from our live smart-phone application request data is currently around 29%.

The offline speech transcription technology is the basis for three applications developed in recent years: a web-based speech transcription service, *Transcribed Speech Archive Browser*, and *Diktofon* – a transcribing voice recorder for the Android smartphone platform. The web-based speech transcription service is a simple front end to the offline speech transcription

technology. It allows users to upload a speech recording, transcribes the recording and sends the transcription to user's e-mail. *Transcribed Speech Archive Browser* is a web interface for accessing large transcribed spoken data collections (in our case, talk shows from several Estonian radio stations, updated daily). The system uses automatic time-aligned transcriptions with speaker segmentation information to present structured speech data more efficiently and make accessing relevant speech data quicker. The Android application *Diktofon* is a voice recorder that can automatically upload recordings to our transcription server, fetch the transcriptions and present them to the user in sync with the recordings.

Our most recent application *Kõnele* is an Estonian real-time speech recognition module for the Android platform. Android comes with default Google-provided speech recognition capabilities for many languages. As Estonian (and also Finnish) is not yet in the set of supported languages, we developed a module that provides Estonian speech-to-text via our public online recognition server. The module implements Android open speech recognition APIs and thus can be used in all applications that call a speech recognition service (such as many keyboard, SMS, navigation applications). Many noteworthy tricks were implemented to make speech recognition more robust and responsive. On the recognition server, unsupervised adaptation was applied that reduced recognition errors by almost a half. The application has become very popular among Estonian Android users.

Recent developments of statistical parametric speech synthesis system GlottHMM

Tuomo Raitio1, Antti Suni2, Martti Vainio2, Paavo Alku1

1Department of Signal Processing and Acoustics, Aalto University, Finland

2Institute of Behavioural Sciences (Phonetics and Speech Synthesis Research Group), University of Helsinki, Finland

This paper describes the recent developments of the statistical parametric speech synthesis system GlottHMM. GlottHMM is a collaborative project between Aalto University and University of Helsinki to create a high-quality speech synthesis system. GlottHMM uses a physiologically oriented modeling approach where speech is decomposed into the glottal source signal and the model of the vocal tract filter through glottal inverse filtering. This enables the individual and detailed modeling and control of the both speech components.

However, the voice source singal is very difficult to model accurately and thus a library of glottal source pulses is extracted from the estimated voice source signal in the analysis stage. In synthesis, a realistic glottal excitation is reconstructed by selecting appropriate glottal source pulses from the pulse library according to the source features. This procedure greatly improves the naturalness of the synthetic speech and its similarity to the original, and it has a great potential to model various phonation and speaking styles.

The modeling of high-pitched female voices has also been difficult with conventional tools such as linear prediction. In GlottHMM system, this has been addressed by using more appropriate techniques such as stabilized weighted linear prediction or pitch-synchronous spectral analysis, which yield better spectral estimates and thus enhance the quality of the synthetic speech.

Vulnerability of speaker verification systems against voice conversion spoofing attacks

Tomi Kinnunen, Zhi-Zheng Wu, Kong Aik Lee, Eng Siong Chng, Haizhou Li

Voice conversion – the methodology of automatically converting one's utterances to sound as if spoken by another speaker – presents a threat for applications relying on speaker verification. We study vulnerability of text-independent speaker verification systems against voice conversion attacks using telephone speech. We implemented two voice conversion systems with two types of nonparallel frame alignment methods and five speaker verification systems ranging from simple Gaussian mixture models (GMMs) to state-of-the-art joint factor analysis (JFA) recognizer. Experiments on a subset of NIST 2006 SRE corpus indicate that the JFA method is most resilient against conversion attacks. However, even it experiences more than 5-fold increase in the false acceptance rate from 3.24 % to 17.33 %.

A sign annotation tool for phoneticians

Stina Ojala1,* and Pertti Palo2,3

1Bioinformatics, Department of Information Technology, University of Turku

2Institute of Behavioural Sciences, Faculty of Behavioural Sciences, University of Helsinki

3Institute of Mathematics, School of Science, Aalto University

*Corresponding author: stiroj@ utu. Fi

When phoneticians want to analyze acoustic speech – i.e., speech – there is a standard system to use: Praat. But when phoneticians want to analyze sign language – i.e., signed speech – there are no adequate systems to choose from. Without such a system phonetic research of sign language is very slow if not outright impossible.

An adequate system should consist of at least a video annotation tool, articulatory models – i.e., models for both hands and facial expressions, and relevant statistical tools such as time series analysis etc. The video annotation tool should – in a mature version – be capable of semiautomatic or even fully automatic tracking of all the articulators.

In this paper, we present a limited prototype of a video annotation tool and a pilot video analysis carried out with it.

Factors affecting the categorizing and rating the goodness of L2 vowels

Terhi Peltola

University of Helsinki

ELTE, Budapest

Foreign accent sometimes causes difficulties for the listener to understand the language learners' speech. Especially, when the correct pronunciation of the foreign speech sounds is problematic for the learner due to category goodness correspondence between speech sounds (Best 1991). For Hungarian Finnish learners the most problematic Finnish vowels are /æ/ and /e/, due to phonemic and orthographic differences. This can sometimes create both confusions and funny sentences, such as Hän lehti takaisin instead of Hän lähti takaisin ('she leaf back' instead of 'she went back').

The current paper is an on-going quantitative investigation on which factors affect the categorization and goodness rating of foreign pronounced vowels. The stimuli were extracted from one previous recordings of a study, in which the effect of means of producing the problematic vowels was investigated (Peltola 2011). The different means of production were reading and imitating. In the present study Finnish university students rated the goodness of these problematic vowels pronounced by Hungarian students separately and in simple CV-syllables /kV, pV, tV/ on the Likert scale (from 1–7).

Three hypotheses are tested in the current paper. Firstly, the effect of the means of production of the speech is investigated: are the L2 read and imitated vowels categorized and rated differently by native speakers? Secondly, the effect of musicality of both the speaker and the rater are investigated. Thirdly the effect of context is investigated. The vowels and the syllables were extracted from the same words.

More detailed description of the results will be given during the XXVII Fonetiikan päivät in Tallinn, Estonia.

References

Best Catherine T. 1991: The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model. Haskins Laboratories status report July-December 1991. SR·107/ 108, 1–30. Haskins Labs., New Haven, CT.

Peltola Terhi 2011: Some remarks on the effect of imitation in novel vowel qualities. In Folia Uralica Debreceniensia 17. 46–53. Debreceni Egyetem finnugornyelvtudomány tanszék, Debrecen.

Ikääntyminen ja vieraan kielen havaitseminen

Henna Tamminen1,2,3, Maija S. Peltola1,2,3, Katri Jähi1,3, Teija Kujala4,5, Heikki Hämäläinen2,6 & Risto Näätänen5,7,8

1Fonetiikka, Turun yliopisto
2Kognitiivisen neurotieteen tutkimuskeskus, Turun yliopisto
3LAB-lab, Fonetiikka, Turun yliopisto
4Cicero Learning Network, Helsingin yliopisto
5Kognitiivisen aivotutkimuksen yksikkö, Helsingin yliopisto
6Psykologia, Turun Yliopisto
7Center of Functionally Integrative Neuroscience, University of Århus, Denmark
8Department of Psychology, University of Tartu, Estonia

Ihmisen kognitiiviset kyvyt hidastuvat monilta osin ikääntymisen myötä. Puheen ja kielen prosessointi kuuluvat näihin kognitiivisiin toimintoihin, joten voidaankin olettaa, että ikääntyneiden ihmisten hidastunut kielellinen prosessointi tulee esille myös vieraan kielen oppimisessa. Tämän tutkimuksen tarkoituksena oli selvittää, miten artikulatorinen kuuntele–ja–toista -harjoittelu vaikuttaa vieraan kielen oppimiseen ikääntyneillä ihmisillä.

Tutkimus oli kolmipäiväinen. Koehenkilöt osallistuivat jokaisena päivänä identifikaatio- (ID), hyvyydenarviointi- (GR), diskriminaatio- (DISK) ja reaktioaikakokeeseen (RT) sekä kuuntele–ja–toista -harjoitteluun. Lisäksi tehtiin herätevastemittaukset, joissa kiinnostuksen kohteena olivat poikkeavuusnegatiivisuusvasteet (mismatch nagativity, MMN). Keskimmäiseen koepäivään kuului kaksi harjoittelua, joten yhteensä harjoittelu tehtiin neljä kertaa. Kokeessa käytetyt ärsykkeet olivat englannin kielen sanat /fi:l/ ’feel’ ja /vi:l/ ’veal’, jotka eroavat toisistaan ensimmäisen äänteen soinnin alkamisajan suhteen. Sanan /fi:l/ ensimmäinen äänne /f/ on suomenkieliselle tuttu, mutta sanan /vi:l/ /v/ on vieras äänne, joka monesti sekoitetaan suomen kielen äänteeseen /ʋ/. Englannin /f/ ja /v/ ovat samankaltaiset, minkä vuoksi niitä on vaikea erottaa toisistaan ja tällöin myös oppiminen on haasteellista. Sanoista syntetisoitiin 15 ärsykkeen jatkumo siten, että ensimmäinen ärsyke oli täysin soinniton ja soinnin alkamisajankohtaa aikaistettiin 14 ms välein viimeisen ärsykkeen ollessa täysin soinnillinen. Kaikki ärsykkeet esiintyivät ID- ja GR-kokeissa ja niiden avulla saatiin selville kategoriarajan jyrkkyys ja sijainti sekä harjoittelun aikaansaamat muutokset. Lisäksi voitiin tarkastella prototyyppien mahdollisia muutoksia. DISK- ja RT-kokeissa, sekä MMN-mittauksissa käytettiin oddball-paradigmaa, jossa käytetyt ärsykkeet olivat lähellä toisiaan, mutta eri kategorioissa äidinkielenään englantia puhuvilla, aikaisemmin testatuilla henkilöillä. DISK- ja RT-kokeilla saatiin selville reaktioajat ja erotteluherkkyydet, sekä niissä tapahtuneet muutokset. Koehenkilöinä oli yksitoista 61–71-vuotiasta vähintään vuoden poissa työelämästä ollutta henkilöä.

Tulokset osoittivat, että koehenkilöt eivät oppineet havaitsemaan uutta äännettä. Harjoittelu sai aikaiseksi joitakin behavioraalisia muutoksia, mutta muutokset eivät viittaa oppimiseen. Kategoriaraja siirtyi, mutta rajan jyrkkyys, reaktioajat ja erotteluherkkyydet eivät muuttuneet harjoittelun myötä. MMN-mittausten tulokset ovat vielä analysoimatta, mutta esittelemme ne Fonetiikan päivillä helmikuussa. Ikääntyminen siis vaikuttaa vieraan kielen äänteiden oppimiseen, sillä samalla koeasetelmalla on saatu selkeitä oppimistuloksia nuorilta aikuisilta.

Ääntämisharjoitusten vaikutus vieraan kielen oppimiseen

Kimmo Peltola, Antti Saloranta, Henna Tamminen, ja Maija S. Peltola

Fonetiikka, Turun yliopisto

LAB-lab, Fonetiikka, Turun yliopisto

Vieraan kielen oppimisessa on yhtenä keskeisenä tutkimuskohteena ääntämisen autenttisuuden saavuttaminen. Yhtenä mahdollisena keinona on kielistudioharjoittelu, jonka perustana toimii kuuntele ja toista –harjoite. Tämä laajasti kouluissakin käytetty harjoituskeino on opetustyössä hyväksi havaittu, mutta tieteellistä näyttöä menetelmän toimivuudesta suhteessa erilaisiin harjoitteisiin on hyvin vähän. Laboratoriomme tutkimustyö alkaa harjoitemenetelmien toimivuuden arvioinnilla, minkä jälkeen voimme tutkia erilaisten harjoitteiden vaikutuksia eri taustaisilla oppijoilla.

Oppimisteorioiden (Flege, 1987, Speech Learning Model; Best & Strange, 1992, Perceptual Assimilation Model) mukaan vaikeimpia opittavia äänteitä ovat sellaiset äänteet, jotka muistuttavat äidinkielen äänteitä, mutta eroavat joko foneettisilta yksityiskohdiltaan (seurauksena aksentti) tai fonologisilta rooleiltaan (fonologinen virhe) äidinkielessä ja kohdekielessä. Koko tutkimuskartoituksen lähtökohtana ovat puolisynteettiset pseudosanaärsykkeet, joissa esiintyy oppimisteorioiden valossa hankalaksi luokiteltava äännekontrasti. Sanaparin ensimmäisessä sanassa [ty:ti] on suomalaisille tuttu pyöreä suppea etuvokaali, kun taas toisessa sanassa [t□:ti] esiintyy erityisen haasteellinen pyöreä keskinen vokaali, joka suomessa assimiloituu lähinnä kategoriaan /y/.

Tutkimusprotokolla on perusteeltaan samanlainen kaikissa harjoitetyypeissä. Tutkimussarjan aluksi tehdään lähtötasomittaus, jossa koehenkilö ääntää molemmat sanat kymmenen kertaa mallin mukaan. Tämän jälkeen vuorossa on harjoitusosio, jossa molemmat sanat toistuvat 30 kertaa. Sitten vuorossa on jälleen mittaussessio ja vielä harjoittelu. Seuraava päivä käynnistyy harjoittelulla, jota seuraa mittaus, sitten jälleen harjoittelu ja lopuksi vielä viimeinen mittaus. Tällöin mittauskertoja on yhteensä neljä ja jokaisesta produktiosta mitataan kohdevokaalin ensimmäinen, toinen, ja kolmas formantti sekä perustaajuus. Kartoituksessamme on viisi harjoitetyyppiä. Ensimmäinen harjoite on *Passiivinen kuunteleminen*, jossa koehenkilö ei keskitä erityistä huomiota kuulemiinsa ärsykkeisiin. Toisessa *Aktiivinen kuunteleminen* -harjoitteessa erona on se, että koehenkilön huomio kiinnitetään kuunneltaviin sanoihin (laskeminen, napin painallukset). Kolmas tyyppi on perinteinen *Kuuntele ja toista*, jossa koehenkilö toistaa kuulemansa sanat. Neljännessä tilanteessa *Foneettiset ohjeet* koehenkilöille annetaan lisäksi ääntämisohjeita ja palautetta tasaisin väliajoin. Viides harjoitetyyppi on *Audiovisuaaliset ohjeet*, joka jakautuu kahteen visuaalisten vihjeiden tyyppiin, sillä puolet koehenkilöistä saa ääntämisohjeet foneettisena transkriptiona ja puolet ortografisina merkkeinä.

Kartoituksen tarkoituksena on selvittää, miten erilaiset harjoitemuodot auttavat autenttisen ääntämisen oppimisessa. Saavutettujen tulosten perusteella jatkamme erilaisille oppijoille

optimaalisten harjoitteiden etsimistä ja yhdistämme tutkimuksiin myös behavioraaliset havaitsemiskokeet sekä MMN-rekisteröinnit.

Kiitokset yhteistyökumppaneillemme Sanako Oy, prof. Teija Kujala, prof. Risto Näätänen, prof. Paavo Alku, prof. Valeri Hazan, sekä opiskelijoillemme Katri Jähi, Oskari Tulonen, Henriikka Uusitupa, Elisa Reunanen, Elina Mehtälä, Arvo Mäkilä, Teijo Lehtinen, Inna Demidova, Johanna Ala-Karvia.

Emotions in cents: A preliminary study of the musical intervals in emotional speech samples

Teija Waaramaa[1], Mona Lehtinen[2]

[1]Speech and Voice Research Laboratory, University of Tampere, Tampere, Finland; teija.waaramaa@uta.fi

[2] Institute of Behavioural Sciences Speech Sciences, University of Helsinki, Helsinki, Finland

Rationale

Fundamental frequency (F0) and sound pressure level (SPL) have been found to be crucial factors in communicating emotions, and these parameters are also much studied. However, the pitch variation in speech, its musical interval analysis and its ability to convey emotional content have been much less studied.

Objective

The aim of the present preliminary study was to investigate whether listeners can perceive emotions from vocally expressed nonsense sentences, and whether the musical intervals measured in cents (100 cents = ~ one semitone of a musical scale) affect the interpretation of the emotion perceived.

Materials and methods

Nonsense sentences (dur ~ 10 s/sentence, N 8) were produced by professional actors (n 2) and actresses (n 2). Emotional states of joy, surprise, interest, sadness, fear, anger, disgust and a neutral emotional state were expressed. The recordings were made by Sony Sound Forge 9.0 recording and editing system and Rode NTK microphone in a professional recording studio. Speakers' distance from the microphone was 40 cm. The acoustic analyses and pitch variations were studied with Prosogram-script for Praat softwere (1) developed by Piet Miertens, 2002 (2). The cents were calculated using the average F0 values of the syllable in question and the following steady one, and converting the values into cents. In the listening test the samples were replayed to 7 male and 7 female listeners. The listeners' answers and gender differences were statistically investigated by PASW18.

Results

The pitch variations in the cents tended to be extremely wide for the two-sentence samples studied, which made the program impossible to find any coherence in them. Several intervals were found in each emotional state. Prosogram could not identify all the syllables in each sample in the measurements. Thus the number of the syllables varied between the emotion samples for the results. The range of the cents measured from the steady state syllables was calculated, and the order from the lowest to highest was the follows: fear, sadness, surprise, anger, neutral, interest, joy and disgust. Correspondingly, the order of the mean from the lowest to highest: neutral, disgust, sadness, interest, fear, anger, joy and surprise. In the listening test, the percentage for the recognition of the emotion samples was 82 %. Most

frequently interest was chosen as an answer, disgust most rarely. Surprise was confused with expressions of fear and interest. Some listeners confused joy and interest with surprise, and fear with sadness. However, these confusions were not statistically significant. The accuracy in the perception of anger was 100 %. No gender differences were found.

Discussion and conclusion

As the duration of each emotion sample was long enough to carry several intervals, no clear pitch contour could be connected to an emotional state. Thus, the sentences studied for cents should be shorter with less variation in the pitch contour. The negative emotional states of fear and sadness were confused with each other obviously for their low arousal level and similar pitch range (cents). The positive emotions of joy and surprise also shared a similar pitch range and were confused with each other. Motivational and survival-related aspects from the evolutionary viewpoint may explain why interest and anger were both prominent among the answers.

References

(1) Boersma P & Weenink D. (2005). Praat: doing phonetics by computer (Version 4.3.01) [computer program]. Retrieved from http://www.praat.org/. (2) Miertens P. (2002). Prosogram (version 2.85) [computer program]. http://bach.arts.kuleuven.be/pmertens/prosogram/.

Emotsioonid ja kõne ajaline struktuur

Kairi Tamuri, Meelis Mihkla

Eesti Keele Instituut

Emotsionaalse kõne uurimused väidavad, et igal emotsioonil on just talle omased akustilised parameetrid, mis teda teistest emotsioonidest ja neutraalsest kõnest eristavad. Üheks selliseks tunnuseks peetakse kõneüksuse kestust. Uurimistulemused näitavad, et võrreldes neutraalse kõnega on rõõmu- ja vihalausungite tempo kiirem ja kurbuselausungite tempo aeglasem. (Murray ja Arnott 2008, Braun ja Oba 2007, Yildirim jt 2004, Iida jt 2003, Banse ja Scherer 1996)

Antud ettekanne keskendub küsimusele, kas ja kuidas avaldavad emotsioonid mõju eestikeelse etteloetud kõne ajalisele struktuurile. Vastuse saavad kaks uurimisküsimust: a) kas emotsioonid mõjutavad kõnetempot ning b) kas emotsioonid tekitavad erinevusi sõnaprosoodias. Uurimuse akustiline baas oli Eesti Keele Instituudi eesti emotsionaalse kõne korpus. Uurimismaterjali hulka valiti sealt välja laused, mille emotsiooni või neutraalsust on tajunud üle 50% kuulajaist ning milles emotsiooni kannab edasi ainult hääl.

Uurimisküsimustele vastuse leidmiseks mõõdeti esmalt emotsioonilausungite artikuleerimiskiirust (häälikute arv sekundis) ning võrreldi saadud tulemusi neutraalse kõnega. Tulemustest selgus, et emotsionaalse ja neutraalse kõne tempod on erinevad ning need erinevused on olulised ka statistiliselt. Teisena uuriti vokaalikeskse struktuurimalli [C]V[V]CV ajaliste tunnuste seoseid emotsioonidega, arvestades sõna rõhulisust, asendit fraasis ja sõnaliiki. Mõõdeti häälikukestusi ja arvutati kestussuhteid V1/V2, V1/C1 ja V2/C2. Analüüsi käigus kombineeriti erinevaid tunnuseid, kuid tulemused näitasid, et seosed on marginaalsed ning sõnaprosoodia ajalises struktuuris emotsioon esile ei tule.

KIRJANDUS

Banse, Rainer; Scherer, Klaus R. 1996. Acoustic profiles in vocal emotion expression. Journal of Personality and Social Psychology, Vol 70(3), 614-636.

Braun, Angelika; Reiko Oba 2007. Speaking Tempo in Emotional Speech – a Cross-Cultural Study Using Dubbed Speech. ParaLing'07, 77–82.

Iida, Akemi; Campbell, Nick; Higuchi, Fumito; Yasumura, Michiaki 2003. A corpus-based speech synthesis system with emotion. Speech Communication, vol. 40, 161–187.

Murray, Iain R.; Arnott, John L. 2008. Applying an analysis of acted vocal emotions to improve the simulation of synthetic speech. Computer Speech and Language, vol. 22, 107–129.

Yildirim, Serdar; Bulut, Murtaza; Lee, Chul Min; Kazemzadeh, Abe; Deng, Zhigang; Lee, Sungbok; Narayanan, Shrikanth; Busso, Carlos 2004. An acoustic study of emotions expressed in speech. INTERSPEECH-2004, 2193–2196.

Pilguheit võru häälikusüsteemi / Lyhykatsaus võron äännesysteemiin

Sulev Iva

Tartu ülikool, Võro Instituut

sulev.iva @ut.ee

Võru keel, suurim vana lõunaeesti keele tänapäevastest järeltulijaist, on oma häälikusüsteemilt väga lähedane nii eesti kui soome keelele. Ühelt poolt on võru keel näiteks oma kindlalt väljakujunenud kolme välte süsteemi ning *õ*- ja palatalisatsioonirohkusega soome keele poolt vaadatuna kindlasti eestilikum ja vähem soomelik kui eesti kirjakeel. Teisalt on aga võru häälikusüsteem mitmes osas, näiteks oma väga hästi säilinud vokaalharmoonia, larüngaalklusiili ja sellega seotud sõnapiiri geminatsiooniga (sm *loppu- tai alkukahdennus*) tunduvalt lähedasem just soome häälikusüsteemile. Muidugi aga võib nende sarnasuste kõrval eesti ja soome keelega leida võru häälikusüsteemist küllaldaselt sellist, mis eristab seda ühtviisi nii eesti kui soome häälikusüsteemist. Näiteks on võru vokaalharmoonia soome omast oluliselt erinev ja võru palatalisatsioon teiselaadiline kui eesti oma. Ettekandes vaadeldaksegi võru häälikusüsteemi eripärasid kolme lähisugukeele häälikusüsteemide sarnasuste ja erinevuste taustal.

The phonetic and phonological analysis of the fall-rise intonation pattern in the Kihnu variety of Estonian

Eva Liina Asu and Nele Salveste

Institute of Estonian and General Linguistics
University of Tartu

This paper presents an analysis of a fall-rise pattern in the Kihnu dialect of Estonian. Like the dialect of the Saaremaa its intonation has been singled out for a peculiar speech melody, which has often been attributed to a possible influence from an earlier substratum of Swedish (Nurmekund 1937). The characteristic quality of Kihnu speech melody has been claimed to be disappearing in the speech of younger islanders and in particular men (Saar 1934).

In the Saaremaa variety, the impression of 'singing' intonation is phonetically given by systematically later peak alignment as compared to Standard Estonian (Asu 2005), whereas the Kihnu variety exhibits frequent rising patterns (Tanning 1948). At closer examination these appear to be fall-rise accents which occur mostly in phrase-final position. Such accents have not been accounted for in the intonational inventory of Standard Estonian (Asu 2004).

The aim of the present study was to investigate in detail the phonetic realisation and segmental anchoring of the fall-rise pattern, and on the basis of this propose the most optimal phonological analysis. Recordings of spontaneous speech from three generations of speakers (both men and women) were used: (1) year of birth around 1900, (2) around 1940, and (3) after 1970. All instances of fall-rise accents were extracted and analysed. The position of the high target relative to the onset of the stressed syllable, the location of the low target and the span of the initial fall as well as the final rise at the intonational boundary were measured. The analysis of the fall-rise accent of Kihnu as H*L H% (within the autosegmental-metrical theory of intonation) was based on the fact that the domain of the intonation pattern varied from a monosyllabic word to utterances of several syllables. This implies that the pattern is not composed of different tunes but constitutes one pitch accent.

References

Asu, E. L. 2004. The Phonetics and Phonology of Estonian Intonation. Doctoral dissertation. University of Cambridge.

Asu, E. L. 2005. Intonational contour alignment in Saaremaa and Standard Estonian. *Linguistica Uralica XLI 2005, 2*, 107-112.

Nurmekund, P. 1937. Om Kihnu. Svio-Estonica. Tartu.

Saar, T. 1934. Murdeülevaade. EKI murdearhiiv.

Tanning, S. 1948. Murdepäevik. EKI murdearhiiv.

Revisiting the Meadow Mari vocalic system

Dennis Estill

University of Helsinki
Department of Finnish, Finno-Ugrian and Scandinavian Studies

Meadow Mari is a Finnic language spoken in certain parts of Russia, especially in the region of the Volga Bend. There are two literary Mariic languages, Meadow and Hill. The purpose of this article is to define the present system of vowels in Standard Meadow Mari on the basis of empirical research and in the light of traditional concepts. Traditional descriptions of the vocalic system vary considerably both on the close/open and forward/backward scales (compare Alhoniemi, Bereczki, Grigoryev, Pengitov). The specific objectives of this experiment were (1) to ascertain whether the three so-called weak full vowels can be described as such or whether they are better described as reduced vowels, (2) to determine the quality of the central vowel, and (3) to measure the extent of roundedness of all vowels. The paper includes a brief consideration of personal features that may be apparent from analyses of formants 3 and 4. Although formants 3 and 4 provide information about roundedness, their wide variation between individuals offers clues concerning speaker-specific vowel quality. In considering problems (1) and (2) it is necessary to calculate those acoustic parameters that bear on reduction and the integrity of the individual sound of the respective vowels.

This was analyzed from four perspectives: (a) the positions of the vowels on a chart displaying formants 1 and 2, (b) vowel length or duration, (c) pitch or fundamental frequency and (d) loudness or intensity. In attempting to resolve the problem referred to in (3) above formants 2, 3 and 4 were analyzed, although less emphasis was placed on formant 2 for reasons mentioned in the paper. The article presents the results of a comprehensive study of 2,274 Meadow Mari vowel tokens from a corpus of approx. 500 words. The method used is statistical analysis of acoustic measurements using a speech analysis computer program. The findings, which are based on evidence from two informants, suggest that there are only two vowel categories in Meadow Mari, viz. full and reduced, as opposed to views which add a third category, weak. Further, the central vowel is reduced and it is neutral with regard to roundedness, presumably depending on the environment. Roundedness can only be viewed as occurring on a continuum in the case of the vowels /a/, /e/, /u/, /ö/ and /ı/ and not as categorical. Attention is further drawn to the need to determine which symbols should be used to describe the central vowel. Generally speaking IPA ɣ and SUT *ə* have been used in this connection. ɣ seems unsatisfactory because it implies unroundedness and should probably be replaced by the roundedness-neutral ə. In SUT the central vowel could also very well be represented by the symbol *ə,* although in the vowel chart it occupies the position also held by the rounded, near back and midclose *o_* , that is to say, a central vowel positioned between *o* and *ə*. These findings indicate the need to consider adjusting descriptions of at least some aspects of the Meadow Mari vocalic system.

Regularized all-pole models: Experiments with text-independent speaker verification in noisy environments

Cemal Hanilci, Tomi Kinnunen, Rahim Saeidi, Jouni Pohjalainen, Paavo Alku, Figen Ertas

Regularization of linear prediction based mel-frequency cepstral coefficient (MFCC) extraction in speaker verification is considered. Commonly, MFCCs are extracted from the discrete Fourier transform (DFT) spectrum of speech frames. In this paper, DFT spectrum estimate is replaced with the recently proposed regularized linear prediction (RLP) method. Regularization of temporally weighted variants, weighted LP (WLP) and stabilized WLP (SWLP) which have earlier shown success in speech and speaker recognition, is also introduced. A novel type of double autocorrelation (DAC) lag windowing is also proposed to enhance robustness. Experiments on the NIST 2002 corpus show that regularized LP methods (RLP, RWLP and RSWLP) yield great improvement on recognition accuracy under additive factory and babble noise conditions in terms of both equal error rate (EER) and minimum detection cost function (MinDCF).

Sparse logistic regression fusion for speaker verification

Ville Hautamäki, Kong Aik Lee, Anthony Larcher, Tomi Kinnunen, Bin Ma, Haizhou Li

Fusion of the base classifiers is seen as the way to achieve high performance in state-of-the-art speaker verification systems. Typically, we are looking for base classifiers that would be complementary. We might also be interested in reinforcing good base classifiers by including others that are similar to it. In any case, the final ensemble size is typically small and has to be formed based on some rules of thumb. In this paper, we are interested to find out the subset of classifiers that has a good generalization performance. We approach the problem from the sparse learning point of view. We assume that the true, but unknown, fusion weights are actually sparse. As a practical solution we regularize the weighted logistic regression loss function by the Elastic-Net constraint. Though sparse solutions can be easily obtained using the so-called least absolute shrinkage and selection operator (LASSO), but it does not take into account high correlation between classifiers. Elastic-Net, on the other hand, is a compromise between LASSO and ridge regression constraints. While ridge regression cannot produce sparse solutions, Elastic-Net can. By using sparseness enforcing constraint we are able to improve over the un-regularized solution in all but tel-tel condition.

Comparison of two data collecting systems for sign language articulation studies

Stina Ojala

Department of Information Technology

University of Turku, FINLAND

Einar Meister

Tallinn University of Technology

ESTONIA

The recent development in sign language research has brought forward the need for collecting three-dimensional data in time. The traditional way of recording sign language has been using a single video camera. It is reasonably easy to use and quite easy to obtain. This method, however, lacks the possibility to record data flow in three dimensions and bounds the researcher to using only two dimensional data in time.

Recently the equipment for capturing three dimensional data in time has become more and more available. In this study we tested the NDI motion capture system in Laboratory of Phonetics and Speech Technology at Tallinn University of Technology on how well the equipment was able to capture signed data flow in space and time. We also compared the performance of the traditional single video data and the modern motion capture systems on how well they did measure the parameters of sign language.

Here we present results of the pilot study on the comparative performance data of a single video capture and motion capture system on one single sign language utterance. The comparison was made based on a selection of parameters including non-measurable events percentage and distance measurements between articulators. We also acknowledge the new system in articulatory research, articulograph and the need in the future to be able to get a similar set of data to compare these results with.

New speech corpora at IoC

Einar Meister, Lya Meister, Rainer Metsvahi

Institute of Cybernetics
Tallinn University of Technology, Estonia

The paper will give an overview of the new Estonian speech corpora collected in recent years at the Laboratory of Phonetics and Speech Technology, Institute of Cybernetics at Tallinn University of Technology. The development of these corpora has been funded by the national program *Estonian Language Technology (2006-2010 and 2011-2017)*.

Corpus of Radio News

The corpus includes ca 300 hours of news recordings and more than 8000 pages of digitized news texts. For the transcription of news recordings the Transcriber (http://trans.sourceforge.net) software has been adapted, about 10% of news recordings have been transcribed.

Corpus of talk shows

The corpus includes ca 20 hours of talk shows from different Estonian broadcast companies typically involving a live discussion between the host and 1-2 quests. All recordings are manually transcribed.

Corpus of lecture speech

The corpus includes ca 300 hours of recordings of academic university lectures (33 different speakers) and more than 40 hours of conference presentations (more than 50 different speakers). The presentation recordings of 30 speakers are manually transcribed.

Foreign accent corpus

The corpus consists of laboratory speech recordings read by non-native speakers of Estonian.. Currently about 160 speakers with different language backgrounds have been recorded: Russian (50 speakers), Finnish (30), German (15), French (12), Italian (5), English (4), Lithuanian (3), Spanish (2), Danish (2), Slovak (2), Japanese (2), Swedish (1), Polish (1), Latvian (1), Scottish (1), Irish (1), Azerbaijani (1), Portuguese (1). As a reference 20 native speakers of Estonian have been recorded.

Multimodal corpora

Recently the facilities for the collection of multimodal corpora including articulatory and acoustic data, hand gestures, head movements, mimics, body posture, etc have been set up. In addition to audio & video recorders the facility includes a Laryngograph Processor, an Electropalatograph (EPG), an Electromagnetic Articulograph (EMA), and an infrared 3D Motion Capture System.

The collection of an acoustic-articulatory database using laryngograph, EPG and EMA and viseme database combining video and EMA has been started.

Access to corpora

For corpora management the LAMUS-system (Language Archive Management and Upload System, http://www.lat-mpi.eu/tools/lamus/) has been adapted. All corpora will be made available via the Estonian Language Resources Center (the national project has been launched in 2012).

Finno-Ugric prosody project

Karl Pajusalu

University of Tartu

The paper describes research goals and results of the Finno-Ugric Prosody Project which has been carried out by researchers at the University of Tartu and other institutions since 1999. The initiator and mentor of the project was Ilse Lehiste (1928-2010). The project is aimed at analyzing less-known Finno-Ugric languages using modern experimental tools.

It focuses on the role of suprasegmental features in word structure: stress, quantity and tone, and their phonetic manifestation as prominence, duration, and fundamental frequency. Up to now, three monographs (Lehiste et al. 2003, 2005, 2008) and more than 20 articles have resulted from the project.

The interaction of different suprasegmentals in southern Finnic languages has been a central topic of the project. Besides the Finnic languages, the prosodic structure of the Mordvin and Mari languages has been investigated, there is also a study of Inari Saami in progress. The results of the project indicate that in Finno-Ugric languages various phonetic features are associated with stress and quantity distinctions are often dependent on constraints of the foot structure. In addition, certain segmental effects, e.g. vowel reduction, can be involved in the realization of prosodic contrasts.

References

Lehiste, Ilse; Aasmäe, Niina; Meister, Einar; Pajusalu, Karl; Teras, Pire; Viitso, Tiit-Rein (2003). *Erzya Prosody*. Mémoires de la Société Finno-Ougrienne 245. Helsinki: The Finno-Ugrian Society.

Lehiste, Ilse; Teras, Pire; Help, Toomas; Lippus, Pärtel; Meister, Einar; Pajusalu, Karl; Viitso, Tiit-Rein (2005). *Meadow Mari Prosody.* Linguistica Uralica Supplementary Series, Volume 2. Tallinn: Teaduste Akadeemia Kirjastus.

Lehiste, Ilse; Teras, Pire; Ernštreits, Valts; Lippus, Pärtel; Pajusalu, Karl; Tuisk, Tuuli; Viitso, Tiit-Rein (2008). *Livonian Prosody*. Mémoires de la Société Finno-Ougrienne 255. Helsinki: The Finno-Ugrian Society.

Quantity in Moksha-Mordvin

Niina Aasmäe, Karl Pajusalu, Nele Salveste, Tatiana Zirnask

University of Tartu

This paper will focus on the description of the temporal characteristics of stressed and unstressed vowels in Moksha-Mordvin.

In the introduction, the authors will survey the data available from earlier observations of vowel duration in Moksha mentioned in literature and provide information pertaining to the background of the present research.

Analysis of data including information on the corpus used for measurement and the dialect varieties represented in the recordings of Moksha centres on a comparison of vowel duration data with respect to several phonetic and phonological characteristics of the test words. Among them are: length of the words (from one to six syllables); patterns of stress assignment (stress on initial vs. non-initial syllable, additional stress in words with 3 – 6 syllables); the composition of the syllables within a word.

Ultimately, the results of the analysis of vowel durations in Moksha will be juxtaposed with the relevant data of Erzya obtained by us at a previous stage of the research and the implications of the findings for research into the prosody of the two languages will be formulated.

Comparison shows that vowel duration in Moksha is a more important correlate of stress than in Erzya. The dialects of Erzya have been found to differ, as far as the relation between stressed and unstressed vowel durations is concerned. The dialects that are prototype of the literary language display a tendency towards equal vowel duration in stressed and unstressed syllables, while in the other dialects (including those influenced by Moksha) vowel durations manifest dependence on stress. From the point of view of stress assignment, there is also a difference between the varieties of Moksha and Erzya under consideration – they range between those characterized by the dominance of initial stress (Moksha and a part of the dialects of Erzya) and those, in which stress is relatively mobile (the core dialects of Erzya).

The acoustic characteristics of monophthongs and diphthongs in the Kihnu variety of Estonian

Helen Türk, Eva Liina Asu, Pärtel Lippus, Ellen Niit

Institute of Estonian and General Linguistics, University of Tartu

The aim of this paper is to describe the acoustics of short and long monophthongs and diphthongs in the Kihnu dialect of Estonian. Kihnu variety belongs to the group of Insular dialects, and is spoken on the Islands of Kihnu, Manija and Ruhnu. The vowel system of Kihnu is one of the richest among regional varieties of Estonian. It is, for instance, characterised by persistent vowel harmony and diphthongisation of long vowels. Diphthongisation is more prevalent than in other dialects, although not complete as in certain environments long monophthongs have been preserved (Saar and Valmet 1997). Also, thriphthongs have been observed in certain environments. The vowel system of Kihnu has been described in the accounts of traditional dialectological studies (e.g. Lonn and Niit 2002) but no acoustic analyses have been carried out (except for a small-scale study of short vowels (Türk 2010)).

The data set for the present acoustic analysis comprised of disyllabic test words containing short and long monophthongs and diphthongs embedded in utterance initial and final position of read sentences. The data was recorded by six female subjects aged 23-42 years. In this paper, for the first time for Estonian, the method of formant dynamics was used, i.e. rather than measuring a single point in the middle of the steady state of the vowel equidistant temporal points were used to characterize the formant trajectory within the whole vocalic part. A script was used to calculate the total duration of each vocalic segment and divide into ten equal intervals. For each formant a contour consisting of nine values was obtained. Additionally, in order to track more closely the formant frequency change over the course of vowel's duration, trajectory length (TL) for each separate vowel section was calculated (cf. Jacewicz et al. 2011).

Examination of formant trajectories enabled to characterise the realisation of diphthongs as well as confirm the existence of triphthongs. The characteristics of short and long monophthongs in Kihnu are discussed with reference to earlier results on Standard Estonian.

References

Jacewicz, E., Fox, R.A., Salmons, J. (2011) Vowel change across three age groups of speakers in three regional varieties of American English. Journal of Phonetics 39, 683-693.

Lonn, V., Niit, E. 2002. Saarte murde iseloomustus ja liigendus. In E. Juhkam (ed) Saarte murde tekstid. Eesti murded VII. Tallinn: Eesti Keele Instituut.

Saar, T., Valmet, A. 1997. Kihnu murrakust. In: O. Kõiva, I. Rüütel (eds) Vana kannel VII:1. Kihnu regilaulud. Tallinn: Eesti Kirjandusmuuseum.

Türk, H. 2010. Kihnu murraku vokaalidest. Bakalaureusetöö. Tartu Ülikool.

A preliminary comparison of Estonian and Finnish stops

Kari Suomi

University of Oulu, Finland

Einar Meister

Tallinn University of Technology, Estonia

Leho Võrk (Viron kielen ääntämys, Jyväskylä 1972) suggests that Estonian initial (invariably short) stops are pronounced in a weaker manner than Finnish initial stops, and that Estonian word medial short (Q1) stops are pronounced very "loosely and with a weak air pressure, so that also their explosion burst is weak", implying that this is another difference between the two languages. Intuitively, such differences seem to exist. However, we are not aware of relevant empirical studies.

We have made preliminary analyses of these potential differences by looking at word-initial and medial intervocalic stops in a material that was designed for a comparison of the phonetic realisation of quantity in the two languages. Nine female speakers were recorded in both languages. The materials included, among other consonants, 1078 initial and 785 medial stop tokens.

The target words, closely matched with respect to segmental identity, e.g.

Estonian			Finnish	
Q1	Q2	Q3	C	CC
kade	kate	katte	katu	katto

occurred in prosodically similar carrier sentences.

Using Praat, we looked at the following parameters: occlusion duration, burst duration, and the difference between the mean intensity of the occlusion and the mean intensity of the burst.

We observed several differences between the two languages that can be interpreted to support Võrk's claims, and no differences that argue against his claims. Thus, in the initial position, the duration of the occlusion was shorter in Estonian than in Finnish, and in the medial position, the duration of occlusion of the short stops was similarly shorter in the Estonian Q1 stops than in the Finnish /single/ stops. In the initial position there was no difference in the duration of the explosive burst, but in the medial short stops there was: the burst was shorter in Estonian than in Finnish.

Some apparent burst intensity differences between the two languages were observed, but it became clear that, in order to arrive at any conclusions in this respect, more stringently controlled materials are needed. At the same time, it is possible that the intuitive “weakness” of Estonian short stops, relative to the Finnish ones, has its full explanation in the shorter durations of the occlusion (initial position) and the burst (initial and medial position).

Finally, it must be re-emphasised that this study was very preliminary, and that materials designed for precisely this research question are needed to resolve the issue.

Tavunalkuisten konsonanttiyhtymien ajoitus suomessa

Mikael Hyvärinen, Pauli Härkönen, Michael O'Dell

Tampereen yliopisto

Viimeaikaisten tutkimusten mukaan joissakin kielissä konsonanttiyhtymä tavun alussa asettuu ajallisesti yhtenä yksikkönä suhteessa saman tavun vokaaliin (ns. "C-center effect"), toisissa kielissä jokainen konsonantti asettuu ajallisesti vain suhteessa seuraavaan äänteeseen. Suomessa tavunalkuisia konsonanttiyhtymiä ja niiden ajallisia suhteta saman tavun vokaaliin ei ole juuri tutkittu. Vaikka suomessa konsonanttiyhtymät tavun alussa ovat miltei mahdottomia sanan sisällä, sananalkuisessa asemassa esiintyy nykykielessä verraten paljon konsonanttiyhtymiä.

Tässä tutkimuksessa mitattiin kahdeksan puhujan akustisia kestoja suomen sanoissa, jotka alkoivat /s-/, /t-/, /st-/ ja /ts-/ mutta loppuivat samalla tavalla. Pyrimme selvittämään, onko suomessa C-centerin kaltaista ilmiötä. Lisäksi raportoimme siitä, onko /ts/ tässä suhteessa samanlainen kuin /st/, vai onko esim. merkkejä /ts/:n affrikaattamaisuudesta.

Articulating Finnish vowels: Results from MRI and sound data

Pertti Palo1,2,*, Daniel Aalto1,3, Olli Aaltonen1, Risto-Pekka Happonen4, Jarmo Malinen2, Jani Saunavaara5, and Martti Vainio1

1Institute of Behavioural Sciences, Faculty of Behavioural Sciences, University of Helsinki
2Institute of Mathematics, School of Science, Aalto University
3Department of Signal Processing and Acoustics, Aalto University
4Department of Oral Diseases, Turku Univeristy Hospital and Department of Oral and Maxillofacial Surgery, Univeristy of Turku
5Medical Imaging Centre of Southwest Finland, Turku University Hospital
*Corresponding author: pertti. palo@ helsinki. fi

We present anatomic and acoustic data from a pilot study on the Finnish vowels [A,e, i, o, u, y, æ, ø]. The data was acquired simultaneously with 3D magnetic resonance imaging (MRI) and a custom built sound recording system. The data consists of a single static repetition of each vowel with constant f_0. The imaging sequence was 7.6 s long and had an isotropic voxel size of 1.8 mm. We report qualitative results of listening analysis of audio data and manual analysis of MR images. Our aim is to answer the question "What are the qualities that wholly describe the vowel articulation in this data?"

The main context of this study is the development of a mathematical and computational model of speech production. From this point of view the qualitative approach is motivated by the need to describe the kinds of geometries that the model has to adapt to.

However, in a broader context other aspects will be interesting. From a phonetical point of view it is interesting to see which structures may be considered active and which take a more passive role in producing the vowels. From a linguistic point of view it is interesting to see what the minimal description for the articulations is, and how this description relates to the observed articulations.

We assess the quality of the recorded vowels based on both the audio and the MRI modalities. Both modalities will be subjected to perceptual evaluation in order to estimate the generalizability of the data set. Audio signals are listened by trained phoneticians and any irregularities in comparison with usual Finnish vowel productions such as nasality are reported as well as the representativeness or prototypicality of the vowel samples. Similarly, the MRI data will be perceptually evaluated by comparing it with descriptions of Finnish vowel articulation in the literature. It should be noted that this is the first 3D data set on Finnish and as such is potentially richer in detail than previously collected data. We also measure the smallest opening and its distance from anatomical features for each vowel articulation.

In a preliminary listening test some of the vowels productions were found to have a slight nasal quality. In general, the vowels were found natural and easily categorizable. Preliminary observations from the MRI data include: 1. Some of the articulations are asymmetric. 2. The position of the tongue can be complex. 3. The smallest opening – point of greatest closure – in the vocal tract can be very small. 4. The position of the tongue can be very extreme.

Microduration in Finnish and Estonian vowels revisited: methodological musings

Stefan Werner

University of Eastern Finland

Einar Meister

Tallinn University of Technology, Estonia

The influence of vowel duration on the perception of different vowel qualities in Finnish and Estonian has been the topic of several of our recent studies (e.g., [1]). For the present paper, we reconsidered some of our methodological choices, reanalyzed result data and tried to establish the reliability of our test design.

1. Test-retest reliability: our vowel tests were replicated with some of the original participants.

2. Reanalysis of our binomial response proportion data: instead of only looking at regression coefficients we did a more sophisticated statistical evaluation with generalized linear models.

3. Variation of the test design: the possibility to play the stimuli repeatedly was removed and reaction time measured as an additional response variable.

We will present the new results and our tentative interpretation of them. It is our intention to stimulate a general discussion of methodological issues related to phonetic perception tests.

Virtuaalilapsi oppii puhumaan vanhemman antaman palautteen ja suorittaman matkimisen avulla

Heikki Rasilo, Unto K. Laine, Okko Räsänen

Aalto-yliopiston Sähkötekniikan Korkeakoulu, Signaalinkäsittelyn ja Akustiikan laitos

etunimi.sukunimi(at)aalto.fi

Lasten kielenoppimista lähestytään lapsen ja vanhemman välisen vuorovaikutuksen näkökulmasta. Vanhempana käytetään artikulatorista syntetisaattoria, joka on ohjelmoitu tuottamaan suomen kielessä esiintyviä artikulatorisia konfiguraatioita. Oletuksena on, että vanhempi ymmärtää lapsen puhetta ja osaa suorittaa puheinversion, ts. tietää, kuinka lapsi tuottaa jokeltamansa äänet. Virtuaalilapsi ei tunne kielen foneemeille tyypillisiä artikulatorisia motorisia käskyjä, mutta oppii ne jokeltamalla satunnaisia artikulaatioita, ja parantamalla jokellusta vanhemman palautteen avulla.

Artikulatorinen synteesi optimoi liikeratoja tiettyinä ajanhetkinä määriteltyjen artikulatoristen tavoitteiden (target) välille. Jokaista foneemia kuvaa tietty artikulatorinen motorinen käsky, joka koostuu ennalta määritellyistä lähestymis-, pito- ja vapautusjaksosta. Toisiinsa limittyvät ohjelmat tuottavat koartikulaatioefektin.

Vanhempien on havaittu antavan enemmän positiivista palautetta aikuismaiseen lasten jokellukseen (Gros-Louis et al. 2006). Lasten on myös havaittu parantavan jokellustaan välittömästi sen jälkeen kun vanhemmat ovat antaneet lapselle johdonmukaista palautetta (Goldstein 2003). Virtuaalilapsemme oppii suomenkielen vokaali- ja konsonanttitavoitteet käyttämällä ohjauksessa apunaan ainoastaan vanhemman palautetta. Nämä opittuaan lapsi tuottaa ymmärrettäviä, realistiselta kuulostavia suomenkielen äänteitä ja niiden yhdistelmiä.

Seuraavassa vaiheessa lapsi oppii liittämään vanhemman tuottamat puheäänet omiin äänteisiinsä. Vanhempien on havaittu matkivan lapsiaan huomattavasti enemmän kuin toisin päin (T. Kokkinaki ja G. Kugiumutzakis 2000). Vanhemmat matkivat myös huomattavasti enemmän CV-tyyppistä jokellusta kuin esim. tasaisia vokaaliäänteitä (Gros-Louis et al. 2006). Meidän työssämme lapsi tuottaa kanonista, CVCV-jokellusta, josta aikuinen matkii siinä esiintyvän konsonantin ja vokaalin mielivaltaisessa järjestyksessä.

Lapsi oppii tilastollisesti toisiaan vastaavat piirteet tuotetuissa akustisissa signaaleissa, ja pystyy tämän mallin avulla mappaamaan aikuisen puheen akustiset ominaisuudet oman puheensa akustisiin ominaisuuksiin. Jokelluksen aikana lapsi oppii myös yhteydet omien artikulatoristen tavoitteidensa ja oman akustisen tuotoksen välillä vastaavalla tavalla. Testit osoittavat, että tällaisen interaktiivisen kanssakäymisen jälkeen lapsi oppii matkimaan vanhempaansa vokaalien osalta täydellisesti ja myös konsonanttien osalta merkittävällä tarkkuudella, vastaten lapsen varhaisen vaiheen puheentuottoa.

References

Flash T, & Hogan N (1985) The coordination of arm movements: an experimentally confirmed mathematical model. Journal of Neuroscience, 5(7), 1688-1703.

Gros-Louis, J., West, M. J., Goldstein, M. H., & King, A. P. (2006). Mothers provide differential feedback to infants' prelinguistic sounds, International Journal of Behavioral Development, 30(6), 509–516.

Goldstein, M.H., King, A.P., & West, M.J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. Proceedings of the National Academy of Sciences, 100, 8030–8035.

T. Kokkinaki ja G. Kugiumutzakis. (2000). Basic aspects of vocal imitation in infant–parent interaction during the first 6 months. Journal of Reproductive and Infant Psychology, 18, 3, 173–187.

Uudet muistijäljet syntyvät jo kolmessa päivässä

Henna Tamminen1,2,3, Maija S. Peltola1,2,3, Teija Kujala4,5 & Risto Näätänen5,6,7

1Fonetiikka, Turun yliopisto

2Kognitiivisen neurotieteen tutkimuskeskus, Turun yliopisto

3LAB-lab, Fonetiikka, Turun yliopisto

4Cicero Learning Network, Helsingin yliopisto

5Kognitiivisen aivotutkimuksen yksikkö, Helsingin yliopisto

6Center of Functionally Integrative Neuroscience, University of Århus, Denmark

7Department of Psychology, University of Tartu, Estonia

Tässä tutkimuksessa tarkasteltiin artikulatorisen harjoittelun vaikutuksia vieraan kielen äänteiden oppimiseen. Tavoitteena oli selvittää, miten behavioraalinen ja neuraalinen plastisiteetti mahdollisesti muuttuu harjoittelun myötä. Koeasetelmassa oli lähtökohtana kouluissa laajalti käytetty kuuntele-ja-toista -menetelmä, joten tutkimus antaa tietoa myös tämän opetusmenetelmän toimivuudesta.

Kolmipäiväiseen tutkimukseen osallistui 12 aikuista suomenkielistä koehenkilöä, jotka eivät olleet kielten opiskelijoita. Tutkimuspäivät koostuivat behavioraalisista tietoisen tason kokeista, esitietoinen tason kokeista sekä kuuntele-ja-toista –harjoitteluista. Näiden kolmen peräkkäisen päivän aikana koehenkilöt harjoittelivat neljästi. Identifikaatio- ja hyvyydenarviointikokeissa käytimme ärsykkeinä 15 syntetisoitua sanaa, /fi:l/ ja /vi:l/, jotka erosivat toisistaan ainoastaan ensimmäisen äänteen soinnin alkamishetken suhteen – kokonaan soinnittomasta labiodentaalisesta frikatiivista täysin soinnilliseen. Erottelukokeissa ja mismatch negatiivisuus (MMN) herätevastemittauksissa koehenkilöille soitettiin ärsykkeitä, jotka äidinkielenään englantia puhuvien mielestä edustavat kahta eri kategoriaa.

Harjoittelun avulla voidaan oppia **tietoisesti** havaitsemaan opittavan kielen mukaisesti, sillä harjoittelu näkyy kaikilla behavioraalisilla mittareilla: kategoriaraja siirtyy ja muuttuu systemaattisemmaksi, hyvyysarvio muuttuu, reaktioajat lyhenevät ja diskriminaatioherkkyys kasvaa. Lisäksi harjoittelu synnyttää uusia **muistijälkiä**, mikä näkyy sekä MMN-vasteen amplitudin kasvuna että latenssin aikaistumisena. Tuloksien perusteella on selvää, että nuoret aikuiset vieraan kielen oppijat hyötyvät jo tällaisesta kolmen päivän mittaisesta, pienimuotoisesta harjoittelusta ja että kouluissa käytetty kuuntele-ja-toista –harjoittelu toimii!

Harjoittelu johtaa pysyviin neuraalisiin muutoksiin – uusien muistijälkien juurtuminen vie aikaa

Maija S. Peltola[1,2,3], Henna Tamminen[1,2,3], Teija Kujala[4,5] & Risto Näätänen[5,6,7]

[1]Fonetiikka, Turun yliopisto
[2]Kognitiivisen neurotieteen tutkimuskeskus, Turun yliopisto
[3]LAB-lab, Fonetiikka, Turun yliopisto
[4]Cicero Learning Network, Helsingin yliopisto
[5]Kognitiivisen aivotutkimuksen yksikkö, Helsingin yliopisto
[6]Center of Functionally Integrative Neuroscience, University of Århus, Denmark
[7]Department of Psychology, University of Tartu, Estonia

Uusien neuraalisten muistijälkien syntyminen on edellytys uuden kielen oppimiselle. Useissa tutkimuksissa on MMN-herätevastemittauksien avulla osoitettu, että maahanmuutto, kielikylpy ja jopa luokkahuoneoppiminen voivat aiheuttaa puheäänteiden esitietoisen havaitsemisen muutoksia.

Tämän tutkimuksen tarkoituksena oli selvittää, ovatko Kuuntele ja toista –harjoittelun avulla syntyneet neuraaliset muistijäljet pysyviä ja kuinka vierasta kieltä erinomaisesti osaavien neuraalinen prosessointi mahdollisesti eroaa sellaisten oppijoiden prosessoinnista, joille uusia muistijälkiä on synnytetty nopeasti laboratorio-olosuhteissa.

Tutkimukseen osallistui kaksi koehenkilöryhmää, joista ensimmäisessä oli englannin kielen pääaineopiskelijoita Turun yliopistosta (Ryhmä 1) ja toisessa mahdollisimman yksikielisiä, ei kieliä opiskelevia suomalaisia (Ryhmä 2). Ryhmän 2 koehenkilöt olivat vuotta aiemmin osallistuneet kolmipäiväiseen harjoittelututkimukseen, jossa heille opetettiin oppimisteorioiden kannalta erityisen haasteellista, sointiopposition perusteella eroavaa englannin sanaparia ’feel’ /fi:l/ – ’veal’ /vi:l/. Molemmat ryhmät osallistuivat behavioraalisiin kuuntelukokeisiin (ID, GR, RT, diskriminaatio) sekä oddball MMN-rekisteröinteihin.

Tutkimukset osoittivat, että vuotta aiemmin harjoittelun myötä syntyneet muistijäljet olivat edelleen toimivia. Lisäksi tuloksista näkyi, että normaalilatenssisen MMN-vasteen lisäksi harjoitteluun osallistuneille koehenkilöille oli syntynyt voimakas aikainen negatiivinen vaste, jota ei löytynyt heti harjoittelun päätyttyä. Alustavan analyysin perusteella näyttää myös siltä, että molempien ryhmien aikainen vaste ja myöhäinen vaste ovat keskenään verrannollisia. Behavioraalisia tuloksia ei ole vielä analysoitu, mutta ne tulokset esitellään Fonetiikan päivillä helmikuussa.

Tämän tutkimuksen perusteella voidaan osoittaa, että kielellisellä harjoittelulla syntyneet neuraaliset muistijäljet ovat pysyviä. Tärkeämpänä ja mielenkiintoisempana tuloksena voidaan pitää sitä, että prosessoinnin syvään juurtuminen ei tapahdu heti harjoittelun jälkeen, vaan oppimistulos on havaittavissa vasta myöhemmin. Tätä syvään juurtumista indikoiva aikainen vaste on latenssiltaan lähellä havaintokynnyksen ylittymisestä kertovaa N1-vastetta, jolloin tulos siis viittaisi siihen, että fonologisen prosessoinnin perusta on syvällä havaitsemisprosessoinnin alussa.

Suomen vokaalin kesto venäjää äidinkielenään puhuvilla alkeisoppijoilla

Riikka Ullakonoja
Soveltavan kielentutkimuksen keskus, Jyväskylän yliopisto

Suomen kielessä, toisin kuin venäjässä, vokaalin kesto on merkitystä erottava piirre.Venäjässä kestoa käytetään sen sijaan sanapainon indikaattorina siten, että painolliset vokaalit ovat painottomia pidempiä. Venäjän painollisten ja painottomien vokaalien keston suhde on de Silvan (1999) mukaan sama kuin suomen kielen lyhyiden ja pitkien vokaalien välinen suhde (2:1). Niinpä voisikin olettaa, ettei vokaalin keston tuottaminen suomen kielessä tuota venäjää äidinkielenään puhuvalle oppijalle juurikaan vaikeuksia.

Aiemmat tutkimukset (Ylinen et al. 2005; 2006) venäjää äidinkielenään puhuvien oppijoiden vokaalin keston havaitsemisesta suomen kielessä ovat kuitenkin osoittaneet, että lyhyen ja pitkän vokaalin erottaminen on vaikeaa, jopa niille oppijoille, jotka ovat asuneet Suomessa (ja siten opetelleet suomea) jo pitkään. Vaikka ensimmäiset akustiset tutkimukset syntyperäisten suomenpuhujien kestosta tehtiinkin jo 1970-luvulla (ks. esim. Lehtonen 1970), ei tähän mennessä tietääkseni ole tutkittu tarkemmin keston tuottamista suomen kielen oppijan puheessa.

Tässä esitelmässä keskitytään viiden venäjää äidinkielenään puhuvan 9-12 vuotiaan lapsen lukupuhunnasta mitattuihin suhteellisiin vokaalikestoihin. Lapset ovat syntyneet Venäjällä ja muuttaneet Suomeen viimeisen vuoden aikana. Oppijoiden vokaalikestoja verrattiin viiden suomea äidinkielenään puhuvan samanikäisen lapsen lukupuhunnasta mitattuihin vokaalikestoihin. Aineisto on kerätty osana DIALUKI-tutkimushanketta (www.jyu.fi/dialuki), jossa tutkitaan lukemisen ja kirjoittamisen diagnosointia äidinkielessä ja toisessa/vieraassa kielessä. DIALUKI-hankkeen ensimmäisessä osatutkimuksessa kaikille koehenkilöille tehtiin myös psykolingvistisiä ja kognitiivisia tehtäviä, jotka nauhoitettiin kuulokemikrofoneilla tietokoneelle. Yhden tällaisen tehtävän nauhoitukset ovat tämän esitelmän aineistona. Tehtävässä koehenkilön piti lukea yksittäisiä sanoja niin tarkasti ja nopeasti kuin hän pystyy. Tutkimuskohteena ovat sanojen ensimmäisten tavujen lyhyet ja pitkät vokaalit.

References

de Silva, V. 1999. Quantity and quality as universal and specific features of sound systems: Experimental phonetic research on interaction of Russian and Finnish sound systems. Studia philologica Jyväskyläensia, Jyväskylä: University of Jyväskylä.

Lehtonen, J. 1970. Aspects of quantity in standard Finnish. Studia philologica Jyväskyläensia VI.

Ylinen, S., Shestakova, A., Alku, P. & Huotilainen, M. 2005. The perception of phonological quantity based on durational cues by native speakers, second-language users and nonspeakers of Finnish. Language and Speech 48(3), 313–338.

Ylinen, S., Shestakova, A., Huotilainen, M., Alku, P. & Näätänen, R. 2006. Mismatch negativity (MMN) elicited by changes in phoneme length: A cross-linguistic study.Brain Research 1072(1), 175–185.

Native Finnish and English speakers' fundamental frequency, sound pressure level and long-time average spectrum characteristics in text-reading in Finnish and English

Kati Järvinen

Speech and Voice Research Laboratory, School of Education, University of Tampere, Finland, e-mail kati.jarvinen@uta.fi

Anne-Maria Laukkanen

Speech and Voice Research Laboratory, School of Education, University of Tampere, Finland, email anne-maria.laukkanen@uta.fi

Introduction

The basic hypothesis for this study was that speaking a language other than the native one has an effect on the voice production. Previous studies have shown that the voice may change when speaking a foreign language compared to speaking the native one, and that languages may differ in from each other in terms of vocal characteristics [1-6]. This study aimed to investigate native Finnish and English speakers' fundamental frequency (F0), sound pressure level (SPL) and long-time average spectrum (LTAS) characteristics in foreign language compared to the native language.

Methods

16 native Finnish (8 males and 8 females, mean ages 37,9 and 32,4 respectively) and 16 English (8 males and 8 females, mean ages 42,1 and 32,8 respectively) speakers read a text first in the native language and then in the foreign language. The speech samples (approximately 1 minute in duration) were acoustically analyzed for the mean, standard deviation and range of F0, mean SPL and LTAS characteristics, calculated from the frequency and amplitude of the strongest peaks between 0-1000, 1000-2000, 2000-3000 and 3000-4000 Hz.

Results

The mean of F0 was significantly higher in English than in Finnish for the Finnish males ($p=.012$) and Finnish females ($p= .036$). The range of F0 changed significantly only for the English females ($p=.036$). The amplitude of the peak between 1000 and 2000 Hz changed significantly for Finns ($p=.036$) and for the English subjects ($p=.018$), and the amplitude of the peak between 2000 and 3000 Hz for Finns ($p=.017$) and for English ($p=.027$). No significant changes in SPL were found.

Significant differences between groups were found: in the change of F0 ($p=.013$), in the mean frequency and amplitude of the spectral peak between 1000 and 2000 Hz ($p=.034$ and $p=.011$ respectively) and in the amplitude of the peak between 2000 and 3000 Hz ($p=.002$).

Discussion

The mean F0 of speech changed when speaking a foreign language compared to speaking the native language, and the change seemed to be somewhat related to the target language. The changes in F0 may be a result of adaptation to a certain level the subjects thought the target language is spoken by its native speakers. In LTAS there was a tendency towards a lower frequency in the peak between 1000 and 2000 Hz in Finnish than in English in both groups. The relative amplitude of the peaks between 1000 and 2000 and 2000 and 3000 Hz was lower in English for both groups. Differences in the mean frequency of the spectral peaks may reflect differences in formant frequencies, while differences in the spectral slope may be related to differences in F0 and other culture dependent vocal ideals.

References

[1] Järvinen K, Laukkanen AM, Izdebski K. Voice Fundamental Frequency Changes as a Function of Foreign Languages Familiarity: An Emotional Effect? In: Izdebski K, editor. Emotions in the Human Voice. USA: Plural Publishing. 2008. 1:203-213.

[2] Syrjä T. Vieras kieli suussa. Vieraalla kielellä näyttelemisen ulottuvuuksia näyttelijäopiskelijan äänessä, puheessa jakehossa. (A strange tongue in the mouth. The dimensions of acting in a foreign language in the students' voice, speech and body.) Academic dissertation. Tampere: Tampere University, 2007. (In Finnish, abstract in English)

[3] Ullakonoja R. Da. Eto Vopros! Prosodic Development of Finnish Students' Read-aloud Russian during Study in Russia. Academic dissertation. Jyväskylä Studies in Humanities 151. Jyväskylä: University of Jyväskylä, 2011.

[4] Ohara Y. Performing gender through voice pitch: A Cross-Cultural Analysis of Japanese and American English. In: Pasero U, Braun F, editors. Perceiving and Performing Gender. Opladen/Wiesbanden: Westdeutscher Verlag.1999.105- 116

[5] Ohara,Y. Gender-dependent pitch levels: A Comparative study in Japanese and English. In: Hall K, Bucholtz M, Moonwomon B. editors. Locating power. Proceedings of the second Berkeley Women and Language Conference. Berkeley, CA: Berkeley Women and Language Group 1992. 2: 468-477.

[6] Wagner A & Braun A. 2003. Is voice quality language-dependent? Acoustic analyses based on speakers of three different languages. 15th ICPhS Barcelona. 651-654.

IN MEMORIAM

Matti Karjalainen (02.04.1946 – 30.05.2010)

Professor Matti Karjalainen died on May 30, 2010, at his home in Espoo, Finland, after a prolonged illness. As the eldest of two children, he was born on the 2nd of April in 1946 to a farming family on a small plot of land in the village of Halttula in Hankasalmi. His elementary education began at the local village school while his middle and high-school studies were carried out at nearby Lievestuore from which he graduated in 1965.

Matti began to be interested in electrical motors and radios after finishing elementary school. At the end of the 1950s the first transistors came on the market, and like many other school boys of that time, he gained a new hobby: building pocket radios. The 1950s gave birth to pop and rock music and Matti followed the new hits on Radio Luxembourg as other youth did. In the beginning of the 1960s he received his first acoustic guitar and harmonica. Matti played music himself and was an active listener for a wide range of genres and musical events.

In the autumn of 1965 Matti began his studies at TTKK. The grades that he received on his university entrance examinations would have sufficed for him to enter the Helsinki University of Technology (TKK) but Matti felt that Tampere would make for a more comfortable place of study.

TTKK had just initiated operations as an extension of TKK. One of the responsibilities of the first class of students was to start up student union activities on the campus. The engineering students founded a radio club of which Matti became a member and he also earned his amateur radio technical class license at the same time. For the engineering students he designed powerful sound reproduction equipment during his spare time. By the autumn of 1968 Matti's studies had progressed so far that he obtained his first job as a course assistant at the university. After completing his Master's thesis in 1972 he entered Finnish military service and was stationed at Riihimäki with the Communications Regiment.

1973 was a decisive year for Matti's career. During a weekend in August he and Unto Laine were performing experiments with electronic music in the electronics laboratory of TTKK. While experimenting with analog resonance circuits they produced sounds that were similar to a human voice. They demonstrated this to the department head, professor Boris Segerståhl, who then asked if this could be developed into a real speech synthesizer. After applying for a grant the Finnish Academy funded a small- scale study for them and the spring of 1974 turned out to be the starting point of speech technology research in Finland. That year was filled with intensive development work and by the end of December the developed synthesizer was able to request – with its own voice – for more funding from the Academy!

In 1974 Matti was selected to be the head electronics assistant and during the same year he completed his Licentiate in Technology degree that dealt with power electronics. Two years later in 1976 Matti married Raija Vappu Valtari. A year later their son was born and then two years later their daughter.

Matti Karjalainen successfully defended his doctoral dissertation on speech synthesis in October of 1978. Research into speech synthesis and its applications continued in Tampere up to 1981. During this time the Synte2 and Synte3 speech synthesizers were constructed. Synte2 created much international interest since it was the first portable, unlimited vocabulary, speech synthesizer that had ever been built for any language. What made this technological breakthrough possible was the introduction of microprocessors, which were being applied to different problem domains.

“Speaking machines” created quite a sensation worldwide when they first appeared and they were also covered quite extensively in Finland by newspapers, radio, and television documentaries.

In 1980 Matti and his family moved to Espoo since he had been appointed as an associate professor at TKK’s department of electrical engineering. When Matti started at the acoustics laboratory the available resources were meager: there was only one course assistant’s position (waiting to be filled) and one laboratory technician available – so nearly everything else had to be started from scratch. Little by little research projects began to spring up, resources started to increase, and the number of students attending his courses began to grow.

Under Matti’s direction the development of research specializing in speech technology at Finnish universities started to emerge, and this occurred much earlier, before the mainstream emergence of Finland’s current speech industry. Along with the birth of mobile digital telephone technology in the early 1990s, strong growth in research created new knowhow that was being applied to industry, e.g., for Nokia’s needs. The research in speech technology that Matti ushered in has continued to be actively pursued at the newly formed Aalto University School of Science and Technology.

While working on speech synthesis, Matti noted that no conventional signal analyzer existed that could measure the quality of the generated speech. He realized that this void needed to be filled by developing computational models that were based on the human ear and hearing system. It is algorithms like the standardized PEAQ (Perceptual Evaluation of Audio Quality) and lossy audio coding algorithms, such as MP3, that are based on psychoacoustic models. Matti continued research in psychoacoustics up to the end of his career and applied the principles widely in different areas of sound technology. He published this research and a textbook (in Finnish) that is still in use.

Matti was also interested in the acoustics of musical instruments and the accurate synthesis of their sounds. The acoustics of the guitar was studied and modeled, and along with that, the first scientific investigation into the acoustical properties of the Finnish zither ("kantele") was published. In that paper it was determined what physical properties of the zither create its beautifully bright tonal structure. With these studies a new line of research formed that has expanded over time to cover audio effects and the interaction between humans and machines using the audio channel.

Matti also initiated the modeling of spatial sound, its recording, and research into its perception at the beginning of the 1990s from which two research groups were formed and continue to perform research to this day.

Alongside his research Matti Karjalainen, through his work as a teacher, taught and parented many of the acousticians and sound technologists in Finland. In Otaniemi he introduced the first university level courses in Finland that covered communication acoustics, the programming of signal processors, and speech processing – courses that are still being taught today. Over the decades hundreds of students have journeyed through these courses so that

industry has acquired engineers well versed in sound technology. This has also helped to fulfill the needs of the Finnish mobile phone industry.

While teaching Matti stressed the importance of learning the principles of learning ("learning to learn"). He also found it important that course material cover a sufficiently broad area so that it would form a foundation on which different work scenarios and new research fields could be supported. During his professorship he supervised over 20 doctoral dissertations and 100 M.Sc. theses.

Over the decades Matti Karjalainen's persevering work and sustained interest towards acoustic and psychoacoustic phenomena has borne fruit. Matti as a leader had a heart of gold and let everyone come to him as they were, humbly nurturing and ensuring that every new student could grow and develop.

The genuine enthusiasm that Matti exhibited towards research was infectious and caught on to other researchers. The Acoustics Laboratory grew explosively near the end of the 1990s so that three new professorships in speech and audio signal processing were formed and the number of personnel rose to about 40. The current unit has been integrated into the Department of Signal Processing and Acoustics which is part of Aalto University's Faculty of Electronics, Communications and Automation.

Matti built up a wide base of international contacts. Already in the mid-1970's he began actively participating in international conferences and visited both commercial and academic research institutes. For example, early on he founded good relations with Stanford University that were later strengthened by him spending a sabbatical year in Palo Alto in 1995.

Matti Karjalainen received several awards in recognition of his substantial effort throughout his life towards acoustics and technology related to audio. In 1999 he was promoted to the level of Fellow in the Audio Engineering Society (AES) and in 2006 received their silver medal. For his influential scientific work on perceptually modeling audio signals he was promoted to the level of an IEEE Fellow in December of 2009.

When Matti turned 60 and was already aware of his illness, he established a trust bearing his name whose objective was to support research in acoustics and especially help support gifted young students. Matti's message was that if a student was motivated and gifted then economic reasons should not be a hindrance to pursue a scientific career and make progress within it.

Matti Karjalainen's own life is an outstanding example of this. The fruit of his creativity, diligence, and thinking will continue to be appreciated and enjoyed for a long time in Finnish and international scientific circles, and will be of benefit to his colleagues, new student generations, and all of society.

Matti Karjalainen's urn was quietly laid to rest in Gräsa's cemetery with family present. His life's work can be remembered by a donation to the Matti Karjalainen trust:

Account owner: Akustinen Seura (Acoustical Society of Finland)

Account number: Nordea IBAN FI9810113000205947

Message: Gift to the Matti Karjalainen trust

The research groups in acoustics and sound processing will deeply miss their former leader, friend, and exemplary figure.

http://www.acoustics.hut.fi/notice/matin_muistolle.html

Ilse Lehiste (31.01.1922 – 25.12.2010)

December 25, 2010 died Ilse Lehiste, an outstanding scholar who has made an invaluable contribution to phonetic and phonological studies of Estonian, its cognates and many other languages, as well as to the research of the phonetic realization of the metric structure in the oral poetry of several languages.

Ilse Lehiste was born on January 31, 1922 in Tallinn. Her first school was Lender Gymnasium in Tallinn, the curriculum of which offered, among else, several languages. Of those, especially Latin inspired her early interest in linguistics. Her 12th schoolyear started during World War II, soon after the first Soviet occupation had begun in Estonia.

Upon graduation (true, the gymnasium had been renamed into Secondary School No. 8) Ilse Lehiste spent a year at Tallinn Conservatory studying the piano, but in 1942 she took up Estonian Philology at the University of Tartu. Her studies were discontinued in 1944 as she fled to Germany to escape the second Soviet occupation. After the war she worked as a secretary at the bureau of Baltic University at Pinneberg near Hamburg. Soon she entered the University of Hamburg, taking Germanic philology, Old-Icelandic philology and classical philology (majoring in Latin). In 1948 Ilse Lehiste defended her doctoral thesis at that university and was promoted Doctor of Philology. Her dissertation analyzed the poetry of William Morris, especially those works of the British poet that drew on Icelandic sagas. She continued as lecturer at the same university.

In 1949 Ilse Lehiste left for the USA, where her first jobs included lecturing on German in two small colleges. In a few years she became a student at the University of Michigan, Ann Arbor, where she met their new professor Gordon E. Peterson, one of the founders of acoustic phonetics as a branch of research, who brought along one of the world's first spectrographs. Here Ilse Lehiste enthusiastically experienced the world of experimental phonetics and theoretical linguistics. In parallel with her studies, she published her first articles and worked as an assistant at the communicative sciences laboratory. That is where in 1959 she defended her second doctoral dissertation "An Acoustic-Phonetic Study of Internal Open Juncture" for PhD. The dissertation was published as a monograph supplement to the journal Phonetica. For the next few years she remained with the laboratory as a research associate.

Ilse Lehiste's longest employment was with the Ohio State University, starting from 1963 (professor from 1965, professor emeritus 1987). She joined the Department of Slavic Languages and Literatures where she worked until 1965, when she became chair of the newly founded Department of Linguistics, which position she kept until 1971. She was also acting chair of the same department in 1984-1985 and, once more, head in 1985-1987.

Ilse Lehiste received an honorary doctorate from the Univerities of Essex (1977), Lund (1982), Tartu (1989), and Ohio State University (1999). In 1990 she became a member of the American Academy of Arts and Sciences, in 1998 a foreign member of the Finnish Academy

of Sciences, and in 2008 a foreign member of the Estonian Academy of Sciences. In 2009 her linguistic effort merited the Wiedemann Language Prize.

As a visiting professor Ilse Lehiste delivered lectures in many universities the world over (including Cologne, Vienna, Tokyo, Beijing, Tartu). In that capacity she stayed in Cologne in 1965, in Vienna in 1974 and in Tokyo in 1980.

Ilse Lehiste has said that although she likes lecturing, research has always been nearer to her heart. As for linguistics, Ilse Lehiste spent decades using methods of experimental phonetics to study prosodic issues of many languages (e.g. Estonian and its cognates, Serbo-Croatian, Latvian, English, Japanese), as well as, for example, the acoustics of Estonian diphthongs and consonants, including the palatalization of the latter, and segmental phonetics of several other languages. Researchers of Estonian quantity degrees still keep referring to Ilse Lehiste's article "Segmental and Syllabic Quantity in Estonian" of 1960, as later the quantity-generating durational relations of Estonian syllables (Q1 2/3, Q2 3/2, Q3 2/1) there presented have been proved over and over again both in her own and in other authors' studies of phonetic quantity in read as well as spontaneous Estonian speech.

Ilse Lehiste has authored over 200 articles, over a 100 reviews, and she has been the author, a co-author or editor of 20 books. Her book "Lectures on Language Contact" (1988) testifies to the author being a marvelous lecturer relying on abundant examples from different languages. Another definitive work is "Principles and Methods for Historical Linguistics" (written jointly with Robert J. Jeffers, 1979). Of her earlier works, the more important ones are two books on prosody: "Accent in Serbocroatian" (1963), co-authored by Pavle Ivić, and "Suprasegmentals" (1970), discussing accent, quantity and tone at the time most phoneticians were fascinated by segmental issues. That work still serves as a valuable handbook for prosody researchers.

Prosody has been the focus of several other monographs (and articles) by Ilse Lehiste, be they devoted to spoken language, read poems, or runic songs, e.g. the monograph "Word and Sentence Prosody in Serbocroatian" (1986, co-author Pavle Ivić), "Estonian Prosody: Papers from a Symposium" (1997, co-edited by Jaan Ross). The Estonian publication "Keel kirjanduses" (Language in literature; 2000, compiled by Jaan Ross) contains translated articles part of which also deal with prosodic issues such as metrics and its acoustic correlates in read poems. The joint study of Jaan Ross and Ilse Lehiste titled "Temporal Structure of the Estonian Runic Songs" (2001) surveys Estonian prosody, discussing the metric structure of Estonian folk songs and the mutual influence of language and music.

Personally, I first had the honour to meet Ilse Lehiste in the spring of 1993, when she presented the Psychology Department of the University of Tartu with a speech analyzer (Kay Elemetrics CSL 4300) and held a related special course on Estonian phonetics. That involved fascinating lectures and exciting practical training, which aroused in me a great interest in acoustic phonetics. There must be many others whom Ilse Lehiste has inspired, either as lecturer or supervisor, to take up phonetic studies. My closer contacts with Ilse Lehiste began in 1998, when she initiated Estonian research into the prosody of Finno-Ugric languages. Many students have benefitted from this collaboration, gaining experience in acoustic analysis of speech; this is proved by several Master's theses and one doctoral dissertation finished, while two more doctoral theses are under way. Most of the present results of Finno-Ugric prosody can be found in the following three monographs: "Erzya Prosody" (2003), "Meadow Mari Prosody" (2005) and "Livonian Prosody" (2008). The whole autumn Ilse Lehiste was actively participating in the preparation of a volume on Moksha prosody. All those studies have benefitted from her inspiring, supportive, advising and critical presence, while her criticism was invariably constructive and helpful, never destructive. One could hardly avoid

being impressed by her enthusiasm in linguistic discussions. During the recent years the joint project mentioned took Ilse Lehiste annually to Estonia. Her last visit took place last August before the 11th International Congress for Finno-Ugric Studies held at Piliscsaba, where she participated with an invited talk (jointly with Karl Pajusalu) on the prosodic studies of Finno-Ugric languages.

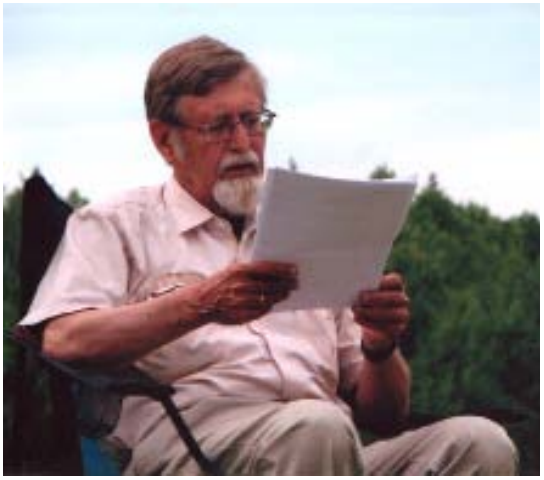
I will always be grateful for the opportunity I had of getting to know such a wonderful person as Ilse Lehiste and of learning from her. Her admirable acumen and intellectual alertness, her integrity and devotion has set an unforgettable example for many of her colleagues and students.

PIRE TERAS (Helsinki—Tartu)

http://www.kirj.ee/public/Linguistica_Uralica/2011/issue_1/ling-2011-1-76-78.pdf

Reprinted with the permission of Linguistica Uralica.

Arvo Eek (13.10.1937 – 09.07.2009)



Arvo Eek was born on October 13, 1937 in Pärnu, a city on the southwestern coast of Estonia. One of five children, Arvo spent his childhood in various places of Pärnumaa County, depending on where his father, a signalman on the railroad, was stationed. After finishing studies at the Eidapere 7-year school in 1952, he continued his education at the Viljandi Pedagogical School which provided vocational training. In 1956, he became a primary school teacher in Luua. One year later, while still teaching, he started long distance studies of Estonian philology at the University of Tartu.

In the autumn of 1959, he became a full-time student and graduated from the University of Tartu in 1963. His last years as a student were a particularly interesting and formative time. In 1961, after a 20 year interlude in Estonian phonetic research, new results based on an instrumental investigation of the Estonian three quantities were published by Ilse Lehiste in the United States and by Georg Liiv in Estonia. In 1962, Georg Liiv defended his Candidate of Philology degree (= PhD) on Estonian vowels in the three quantity degrees. Concurrently, new solutions for Estonian phonology were theorized. Inspired by the results of this research, Arvo Eek became interested in phonetics. Under the supervision of Gerda Laugaste, he wrote his graduation thesis on the articulation and perception of Estonian word quantity, in which he presented acoustic data in a number of oscillograms. He became a graduate student in 1963 at the Institute of Language and Literature of the Estonian Academy of Sciences in Tallinn. His supervisor was Georg Liiv, who had become the head of the phonetics laboratory. However, the laboratory had no equipment. A large analyzer was built for the laboratory at Tallinn Polytechnic Institute, but unfortunately it never functioned. Thus, for a long time, phoneticians had to use the facilities other institutions which owned equipment that could also be applied for phonetic research. It was no wonder that Arvo could not complete his dissertation in time. As a junior researcher of the Institute, Arvo had to work on the side as a proof reader in order to support his family. He was only able to defend his Candidate degree on May 12, 1971.

Arvo Eek's dissertation "Articulation of the Estonian Sonorant Consonants" was written in English. In this work, the articulation of consonants m, n, ń,ŋ, l, l', and r in three distinctive quantities (Q1, Q2, Q3) was studied using static roentgenography, lateral cinefluorography, traditional and direct palatography and filming of lip articulations synchronized with sound spectrography and oscillography. In addition to providing new results on sonorant consonants, the dissertation was the first instrumental study that added to our knowledge of Estonian palatalization (It is the initial, rather than the final part of the consonant that is obligatorily palatalized in an Estonian palatalized consonant — a phonetic characteristic that is inadequately represented by the Uralic Phonetic Alphabet and distorted by the International Phonetic Alphabet). Arvo Eek also observed that the difference between Q2 and Q3 in Estonian sonorant consonants and ŋg is realized in their tenseness (i.e. intensity) rather than in duration.

After receiving his degree, Arvo Eek earned the position of senior researcher and became the head of the phonetics laboratory. Furthermore, under the editorship of Arvo Eek, the series ”Estonian Papers in Phonetics”, published between 1972 and 1987 was successfully launched. From 1972 to 1977, Arvo Eek also taught general and Estonian phonetics at the University of Tartu and supervised several graduation theses in the field of instrumental phonetics. In 1977, the laboratory of experimental phonetics was incorporated into the sector of computational linguistics, which meant certain losses for phonetics. However, this did not stop the development of phonetic research. Arvo Eek carried on working as a senior researcher and a phonetician. In 1987, the Eleventh International Congress of Phonetic Sciences (ICPhS) was organized in Tallinn. During the 1970s and 1980s, Arvo Eek concentrated his attention on the study of various issues related to the Estonian quantity. These included the articulation and perception of duration patterns as well as stress and tone in linguistic feet. In other words, Arvo Eek undertook the most critical problems of Estonian phonetics and phonology. The mounting social reforms and the slow yet noticeable democratization of the Soviet Union during the Gorbaçcev Era gave the first glimmer of hope to the people of the annexed countries for the restoration of independence. Arvo Eek, together with a small group of lawyers and linguists began drafting the initial Language Act. With only minor changes, this document became the first Language Act in the Soviet Union that accorded official status to the indigenous language and consequently limited the scope of the Russian language. Arvo Eek also served as member of the Estonian Congress, the governmental body that adamantly fought for the restoration of the Estonian Republic. From January 1991 to April 1992 he worked part-time as a specialist and until April 1993 he served as an advisor of the State Language Board, dealing with questions related to the regulation of language rights, human rights, and citizenship. His connections with the Institute of Language and Literature became weaker. In 1992, he seized an opportunity to spend six months as a guest researcher at the Laboratory of Phonetics at the Institute of Linguistics at Stockholm University. He left the Language Board in April, 1993 and joined the Laboratory of Phonetics and Speech Technology at the Institute of Cybernetics at the Tallinn University of Technology. From 1993 to 1994, he spent a year doing research in Stockholm. After returning from Stockholm in December 1994, he defended his Doctor of Philology dissertation ”Studies on Quantity and Stress in Estonian” which summarized his work on the subject so far.

His work at the Laboratory of Phonetics and Speech Technology was closely connected with that of Einar Meister, the head of the laboratory and a specialist in speech technology. The two researchers formed a well functioning team, where the ideas and experience of Arvo Eek were combined with Einar Meister’s mathematical competence and knowledge of speech technology. Einar Meister was responsible for generating synthetic stimuli for perception tests, signal processing, statistical analysis, and graphical presentation of experimental data. Together they participated in the international project, ”BABEL – a Multilanguage Database” (1995-1998), creating the Estonian Phonetic Database, which contains all possible diphones and can be used for speech synthesis and analysis. In addition to its practical applications, this proliferant collaboration resulted in 19 scientific publications. In 2003, Arvo Eek together with Einar Meister, Meelis Mihkla, and Heiki-Jaan Kaalep was given the National Science Award for the creation of software for text-to-speech synthesis in Estonian. Arvo Eek continued to work at the Institute of Cybernetics as a senior researcher until the end of 2007.

Shortly after becoming a Candidate of Philology, Arvo Eek aspired to write a comprehensive overview of the phonetics of the Estonian language. Paul Ariste’s work, ”Eesti foneetika” (Estonian Phonetics) (1947) and its second edition ”Eesti keele foneetika” (Phonetics of the Estonian Language) (1953) were based on data gathered in the 1930s and were intrinsically dated. There existed several competing hypotheses about the complicated phonological

structure of Estonian, which meant that new instrumental data were badly needed. With his work, Arvo Eek did his best to fill the existing lacunae. His scholarly style was meticulous; he felt obliged to read everything that was written about Estonian phonetics and phonology, and to survey all viewpoints presented on the topic. The first volume of the much anticipated textbook ”Eesti keele foneetika I” (192 pp) was published at the beginning of 2008. The book provides a thorough overview of Estonian vowels, their articulation and perception, as well as regional characteristics and variation. It also includes an introductory chapter on the role of phonetics and phonology, and articulatory phonetics. Arvo Eek began to prepare the second volume which was to be devoted to Estonian consonants. In 2008, however, he fell seriously ill. Soon after giving a seminar in Tartu only a few days before Christmas, he underwent surgery at the Tartu University clinic. After the operation and a tiresome recovery he made efforts to complete the work on the volume but his time ran out. His colleagues will have to continue his momentous work.

Throughout our acquaintance beginning in our university years, I knew Arvo as an exceptionally balanced person who never raised his voice to anyone. He was, however, deeply insulted by shallow or trite assessments of his carefully formulated results and hard work. His studies in the field of Estonian phonetics provide a solid foundation both for further research and instruction of Estonian phonetics.

TIIT-REIN VIITSO (Tartu)

http://www.kirj.ee/public/Linguistica_Uralica/2009/issue_4/ling-2009-4-313-315.pdf